ARI Research Note 99-31

Expanding the Concept of Quality in Personnel: Final Report

Norman G. Peterson, Lance E. Anderson, and Jennifer L. Crafts
American Institutes for Research

Donald A. Smith
Human Resources Research Organization

Stephan J. Motowidlo
University of Florida

Rodney L. Rosse
American Institutes for Research

Gordon W. Waugh, Rodney McCloy, Douglas H. Reynolds, and Michelle R. Dela Rosa
Human Resources Research Organization

Selection and Assignment Research Unit
Michael G. Rumsey, Chief

June 1999

U.S. Army Research Institute
for the Behavioral and Social Sciences

U.S. Army Research Institute
for the Behavioral and Social Sciences

A Directorate of the U.S. Total Army Personnel Command

EDGAR M. JOHNSON
Director

Research accomplished under contract
for the Department of the Army

American Institutes for Research

Technical review by

Henry Busciglio, Office of Personnel Management

NOTICES

REPORT DOCUMENTATION PAGE

1. REPORT DATE (dd-mm-yy)	2. REPORT TYPE	3. DATES COVERED (from. . . to)
June 1999	Final	15 Sep 92 to 14 Mar 97

4. TITLE AND SUBTITLE	5a. CONTRACT OR GRANT NUMBER
Expanding the Concept of Quality in Personnel: Final Report	MDA903-92-C-0199
	5b. PROGRAM ELEMENT NUMBER 2Q162785

6. AUTHOR(S)	5c. PROJECT NUMBER
Norman G. Peterson, Lance E. Anderson, Jennifer L. Crafts, Donald A. Smith, Stephan J. Motowidlo, Rodney L. Rosse, Gordon W. Waugh, Rodney McCloy, Douglas H. Reynolds, and Michelle R. Dela Rosa	A791
	5d. TASK NUMBER 2211
	5e. WORK UNIT NUMBER C3

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
American Institutes for Research 3333 K St. NW Washington, DC 20007	

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. MONITOR ACRONYM
U.S. Army Research Institute for the Behavioral and Social Sciences Attn: TAPC-ARI-RS 5001 Eisenhower Avenue Alexandria, VA 22333-5600	ARI
	11. MONITOR REPORT NUMBER Research Note 99-31

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT *(Maximum 200 words)*:

This report describes a project designed to determine if new predictors of performance which could increment currently available operational or experimental aptitude measures could be identified. The focus of prediction was performance of army non-commissioned officers (NCO). New tests of NCO situational judgment, prioritization skills, and self efficacy were examined in terms of what they could add in terms of predictive power to a set of available measures--a cognitive composite, a spatial test, and a temperament measure. These were linked to a composite criterion measure based on a structured interview, supervisor behavioral ratings, and supervisor situational ratings.

New and existing measures were administered to 691 non-commissioned officers across four grade levels. The situational judgment test and a situational self-efficacy measure had moderate correlations with a composite criterion and each added a small increment to the validity generated by a combination of the existing predictors. It appears that situational judgment tests and self-efficacy measures have promise in predicting leader performance although most of the predictive variance they provide may be shared with that of tests of cognitive ability and temperament.

15. SUBJECT TERMS

Situational Judgment Test; Self Efficacy; Non Commissioned Officer (NCO); Individual Differences; Assessment of Background and Life Experiences (ABLE); Armed Forces Qualification Test (AFQT); Army; Tacit Knowledge; Predictors; Performance; Ratings; Incremental Validity

SECURITY CLASSIFICATION OF			19. LIMITATION OF ABSTRACT	20. NUMBER OF PAGES	21. RESPONSIBLE PERSON (Name and Telephone Number)
16. REPORT	17. ABSTRACT	18. THIS PAGE			
Unclassified	Unclassified	Unclassified	Unlimited		Michael G. Rumsey (703) 617-8275

Expanding the Concept of Quality in Personnel: Final Report

By

Norman G. Peterson
Lance E. Anderson
Jennifer L. Crafts
Donald A. Smith
Stephan J. Motowidlo
Rodney L. Rosse
Gordon W. Waugh
Rodney McCloy
Douglas H. Reynolds
Michelle R. Dela Rosa

September 30, 1998

Sponsor: *U.S. Army Research Institute for the Behavioral and Social Sciences*
Contractors: *American Institutes for Research*
Human Resources Research Organization

Prepared for U.S. Army Research Institute under Contract No. MDA903-92-C-0199

Executive Summary

The goal of this project was to improve the prediction of performance of NCOs (E-5 through E-8 ranks) by identifying, developing, and validating additional predictor measures hypothesized to increment the prediction of NCO job performance, such as leadership, decision making, and interpersonal abilities. This research followed a concurrent validation design and involved five tasks:

(1) Identification of new predictor constructs,
(2) Identification of performance dimensions,
(3) Development and pilot testing of new predictors,
(4) Development and pilot testing of criterion measures, and
(5) Collection and analysis of concurrent validation data.

During the base year, we conducted the first and second tasks through careful examination of recent research findings and performance models. Our recommendations for predictor measures included new measures that would need to be developed and validated, and marker measures that would provide scores on constructs that are currently in use. These marker measures would provide a baseline of predictive capability to which the new measures could be compared. For the marker tests, we planned to use the ABLE to tap a series of personality and background variables and the ASVAB to serve as markers for cognitive abilities. In addition, we planned on using commercially available instruments independent of the ASVAB to tap cognitive abilities. In terms of new measures, we examined the remaining constructs and determined that most of them fall into the realm of tacit knowledge. In particular, problem understanding, planning, problem solving, practical intelligence, and social intelligence. Further, we posited that tacit knowledge could be tapped by a single, carefully developed, contextually rich instrument. Therefore, we recommended the development of a variant of a written, multiple-choice "situational judgement" kind of test, which we would later call the Army Leadership Questionnaire (ALQ). We also determined that during development of the ALQ, we would attempt to develop other measures which could reflect potentially relevant "single constructs" such as self-efficacy.

In our analysis of various potential criterion measures (Task 2) we considered the costs and benefits associated with developing, administering, and analyzing each of the criterion measures. We recommended collecting a) supervisory ratings (using behaviorally-anchored rating scales developed/modified for NCOs), b) situational interview ratings, and c) self-report data on various administrative actions (using a form based on the Career Force "Personnel File Form" (Campbell & Zook, 1994).

During the second year of the contract, we developed and pilot tested three experimental predictor measures (Task 3). These measures were the ALQ, the self-efficacy portion of the ALQ (SE-ALQ), and the Leadership Problems Inventory (LPI). The purpose of the ALQ was to examine the application of various tacit knowledge constructs to a variety of situationally based problems. For each item, the problem is presented along with five options for responding to the problem. The SE-ALQ is a measure of self-efficacy that requires respondents to indicate, for each item on the ALQ, their degree of confidence that they provided the correct answer to the item. Thus, as each item deals with a different NCO problem, the SE-ALQ might serve as an indication of the self-efficacy of the NCO when actually facing these types of problems. The LPI was designed to tap NCO skills in planning, problem understanding, and problem solving skills. Each LPI item presents the respondent with five problem statements. The respondent is then to indicate the order in which he/she would handle the problems.

To support development of these instruments, we

- gathered problem statements from NCOs,
- edited and refined the problem statements,
- gathered problem solutions from NCOs,
- created an empirically based taxonomy of problems,
- selected problems from the taxonomy for use in item writing,
- wrote items,
- pilot tested two alternate forms,
- gathered keying response data from USASMA NCOs,
- analyzed the pilot test data and keying response data,
- refined the scoring algorithms, and
- created two forms of the instruments for use in the validation study.

One key step in this process was the creation of an empirically-based taxonomy. This taxonomy, based on data gathered from USASMA NCOs, allowed us to categorize problems based on "type of problem" and "relative priority". This taxonomy was important as it provided us with a content-based method for selecting items for use on the ALQ. It also assisted us in development of the LPI, as we were able to select problems for each item that varied according to priority.

Next, we developed two additional measures of self-efficacy. The first measure, Perceived Number of Correct (PNC) asks the respondent to indicate his/her confidence in achieving various levels of performance on the ALQ instrument as a whole. The second measure, Level of Difficulty associated with Leadership Problems (LDLP) provides, for each LDLP item, an NCO problem description. The respondent is then to indicate his/her confidence in dealing effectively with that problem.

Also, during the second year of the contract, we developed and pilot tested the criterion instruments (Task 4). To develop these instruments, we

- developed performance dimensions,
- conducted critical incident workshops,
- revised and finalized the performance dimensions,
- conducted retranslation and instrument development workshops, and
- drafted criterion measures.

In addition to the three criterion measures mentioned above (Personnel File Form, NCO Behaviorally-Anchored Performance Ratings Scales, and the Structured Interview) we developed a fourth measure to gather Situational Performance Ratings.

Finally, we conducted a validation study in which we collected and analyzed criterion and predictor data on nearly 700 NCOs. We created a single criterion score by essentially summing up the instrument scores on the NCO Behaviorally-Anchored Performance Ratings Scales, the Situational Performance Ratings, and the Structured Interview. Then we ran a series of multiple regressions to assess the absolute and incremental validities of the experimental predictor measures over the various marker variables. We corrected for range restriction in ASVAB/AFQT scores with each of these regressions.

We found that the ALQ had a moderate bivariate relationship with the criterion, and it substantially improved prediction of the criterion over the AFQT and other measures of cognitive ability. In addition, the ALQ exhibited no predictive differences between the numerically largest racial subgroups. However, the ALQ added only slightly to the combined predictive power of the AFQT and the ABLE. We found similar results with the LDLP. The LPI did not correlate significantly with the criterion by itself, and failed to increment validity over either the cognitive or ABLE measures.

The pattern of relationships between the NCO performance criteria and the array of predictors indicated that tacit knowledge as measured by the ALQ, self-efficacy as measured by the LDLP, and the ABLE scales of Dominance and Work Orientation contributed most substantially to the effective NCO performance. Cognitive ability, as measured by the ASVAB or similar tests, contributed at a lower level. Interestingly, effective performance earlier in a soldier's career, as indicated by early formal recognition and recommendations for advanced promotion, also showed a substantive relationship with effective NCO performance.

Our major recommendations are:

- Conduct further evaluations of the ALQ and LDLP to determine their operational utility —using a predictive validation design.

- Do not pursue further development or evaluation of the LPI.

- Explore the use of performance measures developed in this project for various applied purposes (e.g., training, coaching, promotion) or in other NCO research.

- Conduct further investigations into the link between performance and administrative self-report indicators such as Recognition, and Recommendations for Accelerated Promotions. Greater understanding of this relationship may contribute to improved prediction of NCO performance.

Abstract

The goal of this project was to improve prediction of performance of NCOs (E-5 through E-8 ranks) by identifying, developing, and validating additional predictor measures hypothesized to increment the prediction of NCO job performance, such as leadership, decision making, and interpersonal abilities. This research followed a concurrent validation design and involved five tasks: (1) identification of new predictor constructs, (2) identification of performance dimensions, (3) development and pilot testing of new predictors, (4) development and pilot testing of criterion measures, and (5) collection and analysis of concurrent validation data.87 Three new experimental predictor measures were developed and pilot tested: (1) the Army Leadership Questionnaire (ALQ), developed to examine the application of various tacit knowledge constructs to a variety of situationally based problems; (2) a measure of self-efficacy; and (3) the Leadership Problems Inventory (LPI), designed to tap NCO skills in planning, problem understanding, and problem solving skills. Our major findings were that the ALQ and self-efficacy measures added incremental validity over measures of cognitive ability, and the LPI did not correlate significantly with the criterion by itself, and failed to increment validity over the cognitive measures. Recommendations were to conduct further evaluations of the ALQ and self-efficacy measures to determine their operational utility--particularly with high demand characteristics situations.

TABLE OF CONTENTS

TABLES

FIGURES

LIST OF APPENDICES

EXPANDING THE CONCEPT OF QUALITY IN PERSONNEL FINAL REPORT

CHAPTER 1

INTRODUCTION

The purpose of this report is to provide a comprehensive summary of the research conducted to develop and validate measures for the Expanding the Concept of Quality in Personnel (ECQUIP) project. This report describes the collection of preliminary data, instrument development, collection of pilot test data, revision of the instruments, collection of the validation data, analyses and discussion of the validity study data, and conclusions.

This report contains seven substantive chapters. Chapter 2 describes the development and evaluation of predictor measures for new and marker tests. Chapter 3 reviews the development and evaluation of criterion measures. Chapter 4 covers the pilot test, including a description of the predictor and criterion measures used and a review of the procedures and results. In Chapter 5, we describe the revisions that were made to the measures in preparation for the validation study, as well as other preparations we made for the validation study data collection. Chapter 6 presents the results of the validation study and provides a discussion of those results. Finally, in Chapter 7, we present our conclusions.

General Background

The general goal of the Project A and Building the Career Force projects sponsored by the Army Research Institute was to examine the predictive validity of the ASVAB (Armed Services Vocational Aptitude Battery) and a set of additional predictors, such as temperament, perceptual, and biographical measures, for soldiers in entry-level Army positions. The ECQUIP project, however, had a different focus. The target population of interest was NCOs. The goal was to improve prediction of performance of NCOs (E-5 through E-8 ranks) by identifying, developing, and validating additional predictor measures hypothesized to increment the prediction of NCO job performance, such as leadership, decision making, and interpersonal abilities. In the process, it was hoped that we would learn more about NCO performance.

Overview of Project

The ECQUIP project had a concurrent validation design with five tasks. During the base year, the first two tasks were conducted. New predictor constructs (Task 1) and performance dimensions (Task 2) were identified through careful examination of recent research findings and performance models (Peterson, Smith, Hoffman, Pulakos, Reynolds, Potts, Oppler, & Whetzel, 1993). Next, we developed and pilot tested measures of the new predictor constructs (Task 3). Then, we developed and pilot tested criterion measures (Task 4) identified during the base year.

Finally, to complete Task 5, we conducted a concurrent validation study in which we collected and analyzed criterion and predictor data on NCOs. Our analyses of these data were meant to determine the absolute and incremental validities of the new predictor measures. The results of these analyses have guided our recommendations about further uses of the measures.

At the end of the base year, as described in the Base Year Final Report (Peterson et al., 1993), we identified a set of constructs and measures that covered the predictor domain. Our recommendations included new measures that would need to be developed and validated, and marker measures that would provide scores on constructs that are currently in use. These marker measures would provide a baseline of predictive capability to which the new measures could be compared. Those recommendations are summarized in Table 1.1.

For the marker tests, we planned to use the ABLE to tap a series of personality and background variables and the ASVAB to serve as markers for cognitive abilities. In addition, we planned on using commercially available instruments independent of the ASVAB to tap cognitive abilities. (We later decided to use cognitive tests developed by the Army and available for use without additional cost.) We felt this was necessary to address difficulties that may be encountered because ASVAB scores may vary in their usefulness for our validation study. For example, the version of the ASVAB taken by the subject may differ, or, at a minimum, the number of years since the scores were obtained would vary greatly.

In terms of new measures, we examined the remaining constructs and determined that most of them fall into the realm of tacit knowledge. Constructs fitting this realm in particular, include problem understanding, planning, problem solving, practical intelligence, and social intelligence. Further, we posited that tacit knowledge could be tapped by a single, carefully developed, contextually rich instrument. Therefore, we recommended the development of a variant of a written, multiple-choice "situational judgment" kind of test, which we would later call the Army Leadership Questionnaire (ALQ). The idea behind the ALQ was to examine the application of these various tacit knowledge constructs to a variety of situationally based problems. The problems would be presented and a response would be obtained in a low fidelity setting (i.e., a paper-and-pencil instrument). The ALQ would have the benefits of examining responses to many situations without the costs associated with more realistic simulations.

We also determined that the ALQ development process, which included gathering information from subject matter experts on a wide range of problem situations, would facilitate development of various "single construct" measures. We planned to evaluate the feasibility of developing these measures in the initial ALQ development workshops. We rank-ordered the constructs from most to least promising, based on likelihood of developing a test for the construct (Peterson et al., 1993, p. 53), as follows:

- Practical Intelligence
- Social Intelligence
- Self-Efficacy
- Declarative/Procedural Knowledge

- Problem Understanding
- Problem Solving/Planning

Table 1.1

Measurement Recommendations from the Base Report

Construct	Measure
Marker Measures	
Cooperativeness	ABLE
Emotional Stability	"
Internal Locus of Control	"
Dominance	"
Achievement Motivation	"
Self-Esteem	"
Dependability	"
Verbal Ability	ASVAB
	Commercial Test
Quantitative Ability	ASVAB
	Commercial Test
Spatial Ability	Assembling Objects, Project A
New Measures	
Enterprising Interests	ABLE
NCO Interests	"
Problem Understanding	Army Leadership Questionnaire
Cognitive Complexity	"
Planning	"
Problem Solving	"
Inference Generation	"
Mental Models	"
Schemes/Scripts	"
Declarative/Procedural Knowledge	"
Practical Intelligence	"
Social Intelligence	"
Self-Efficacy	"

On the criterion side, at the end of the base year, we had explored the possibility of including supervisory ratings, written exercises, role plays, a situational interview, and administrative measures (e.g., reports of awards, training completed, promotion rate, etc.). Based on a careful analysis of the costs and benefits associated with developing, administering, and analyzing each of these criterion measures, we recommended collecting the following:

- supervisory ratings on behaviorally-anchored rating scales developed/modified for NCOs,
- a situational interview to assess 13[1] performance dimensions identified during the base year, and
- a self-report form for recording administrative actions (based on the Career Force "Personnel File Form").

In the base report, we recommended against developing role plays of interactive supervisory tasks; the situational interviews were a less costly alternative, and they had the potential for covering a wider range of situations.

Next, we developed and pilot tested three experimental predictor measures. These measures were the ALQ, the self-efficacy portion of the ALQ (SE-ALQ), and the Leadership Problems Inventory (LPI). In developing the ALQ, we

- gathered problem statements from NCOs,
- edited and refined the problem statements,
- gathered problem solutions from NCOs,
- created an empirically based taxonomy of problems,
- selected problems from the taxonomy for use in item writing,
- wrote items,
- pilot tested two alternate forms,
- gathered keying response data from USASMA NCOs,
- analyzed the pilot test data and keying response data,
- refined the scoring algorithm, and
- created two forms of the ALQ for use in the validation study.

One key step in this process was the creation of an empirically-based taxonomy. We created this taxonomy through a data collection effort undertaken with United States Army Sergeants Major Academy (USASMA) NCOs. As part of this data collection, senior NCOs sorted each problem into categories, and then rated each problem on "relative priority", importance, and frequency. This taxonomy was important as it provided us with a content-based method for selecting items for use on the ALQ. It also assisted us in development of the LPI.

The SE-ALQ is a measure of self-efficacy that requires respondents to indicate, for each item on the ALQ, their degree of confidence that they provided the correct answer to the item. Thus, as each item deals with a different NCO problem, the SE-ALQ serves as an indication of the self-efficacy of the NCO when actually facing many different types of problems. This measure would have the advantage over other potential measures of self-efficacy, in that responses to the item may not be affected by social desirability. Following development of the response format for this measure, we appended that response format to every item on the ALQ.

---

[1] ***The base report identified 15 dimensions. As described in Chapter 3, this was later reduced to 13.***

Then, as the measure was part of the ALQ, we obtained pilot test and keying response data. We used the results of the pilot test to assist us in reducing the number of ALQ items to which the SE-ALQ response format would have to be applied. This allowed us to prepare a more efficient validation study version of the SE-ALQ.

We developed the LPI in response to feedback we obtained from NCOs during development of the ALQ. NCOs told us that the manner in which they decide to deal with a problem is affected by its priority relative to other problems that they must deal with at that time. The base report (Peterson et al., 1993) supported this notion. Several skills that underlie the act of setting priorities were identified as important constructs to measure, including planning, problem understanding, and problem solving. Thus, with the support of Army Research Institute staff, we developed the LPI measure designed to tap NCO skills in planning, problem understanding, and problem solving skills. Each LPI item presents the respondent with five problem statements. The respondent is then to indicate the order in which he/she would handle the problems. In developing the LPI, we

- chose problem statements for use on each LPI item according to the empirically based taxonomy used to select items for the ALQ,
- structured items so that problems varied according to priority as indicated by USASMA NCOs,
- pilot tested two alternate forms,
- gathered keying response data from USASMA NCOs,
- analyzed the pilot test data and keying response data,
- refined the scoring algorithm, and
- created two forms of the LPI for use in the validation study.

Following the pilot test of these experimental predictors, we decided to develop two additional measures of self-efficacy. The first measure, Perceived Number of Correct (PNC) asks the respondent to indicate his/her confidence in achieving various levels of performance on the ALQ instrument as a whole. The second measure, Level of Difficulty associated with Leadership Problems (LDLP) provides, for each LDLP item, an NCO problem description. The respondent is then to indicate his/her confidence in dealing effectively with that problem. Each of these measures used a response format based on the self-efficacy literature (Lee & Bobko, 1994).

Concurrently, we developed the criterion instruments. To develop these instruments, we

- developed performance dimensions,
- conducted critical incident workshops,
- revised and finalized the performance dimensions,
- conducted re-translation and instrument development workshops, and
- drafted criterion measures.

In addition to the three criterion measures mentioned above (Personnel File Form, NCO Behaviorally-Anchored Performance Ratings Scales, and the Structured Interview) we developed a fourth measure to gather supervisor judgments of how the NCO performs in different types of situations. We based this measure on the empirical taxonomy of problems identified earlier in the development of the ALQ taxonomy.

Finally, we conducted a validation study in which we collected and analyzed criterion and predictor data on nearly 700 NCOs. We created a single criterion score by essentially summing up the instrument scores on the NCO Behaviorally-Anchored Performance Ratings Scales, the Situational Performance Ratings, and the Structured Interview. Then we ran a series of multiple regressions to assess the absolute and incremental validities of the experimental predictor measures over the various marker variables. We corrected for range restriction in ASVAB/AFQT scores with each of these regressions and adjusted them for shrinkage.

We found that the ALQ had a moderate bivariate relationship with the criterion, and it substantially improved prediction of the criterion over the AFQT and other measures of cognitive ability. In addition, the ALQ exhibited no predictive differences between racial and ethnic groups. However, the ALQ added only minimally to the predictive power of the AFQT and the ABLE. We found similar results with the LDLP. The LPI did not correlate significantly with the criterion by itself, and failed to increment validity over either the cognitive or ABLE measures.

Our major recommendations are:

- Conduct further evaluations of the ALQ and LDLP to determine their operational utility particularly in "high stakes" situations.

- Do not pursue further development or evaluation of the LPI.

- Explore the use of performance measures developed in this project for various applied purposes (e.g., training, coaching, promotion) or in other NCO research.

- Conduct further investigations into the link between performance and administrative self-report indicators such as Recognition, and Recommendations for Accelerated Promotions. Greater understanding of this relationship may contribute to improved prediction of NCO performance.

CHAPTER 2

DEVELOP AND EVALUATE PREDICTOR MEASURES FOR NEW AND MARKER TESTS

As noted in Chapter 1, after completing work in the base year, we concluded that development of new predictor measures should be pursued, based on a belief that NCO job performance could be predicted by constructs not currently included in the military selection system. Many of the experts who had participated in the initial construct identification phase agreed that developing more complex predictors was a useful step, although operationalizing these measures would be difficult. In particular, many of them articulated the view that the environment, context, situation, or some similar facet was very important in the prediction of NCO performance. Although the experts differed on exactly how to treat these factors, opinion did seem to converge on "building in" to the predictor measurement system some of the environmental richness and complexity represented in NCO jobs. Almost all the experts supported the notion of exploring the development of some sort of complex, simulation-like measure. In addition, for all experts, the measurement operations of choice were complex, interactive, high fidelity simulations that allow derivations of a variety of scores. However, they recognized that such simulations are currently too expensive to develop and implement. Therefore, they suggested a series of compromises. One of the compromises was a more abstract and lower fidelity simulation, much like the Situational Judgment Test developed as part of Project A/Career Force (Campbell & Zook, 1994).

Therefore, a preferred method for obtaining a measure of the constructs recommended in the base report was to develop a written, multiple-choice, situational judgment type of test. This type of test was hypothesized to be useful in tapping NCO Tacit Knowledge. This procedure was also favored to maximize efficiency of assessing those single constructs closely related to tacit knowledge (e.g., practical intelligence, social intelligence) for which no effective measure currently exists for this population. We also decided to attempt to separately measure the predictor construct of self-efficacy with this test. Finally, in the process of developing the situational judgment test, we decided to develop a separate test of examining priority-setting. This leadership behavior also relates to the constructs of problem understanding, problem solving and planning.

A written measure of NCO Tacit Knowledge is a form of job simulation. Job simulations used as predictors can be more or less realistic, depending on the degree of similarity between the simulation and the actual job. Because highly realistic simulations create a close behavioral correspondence between predictor and criterion, they might be better predictors of job performance. But high levels of realism are often prohibitively expensive. Furthermore, if there are predictive gains associated with increasing levels of realism, it is not clear that they are enough to offset the costs of developing and using highly realistic equipment and props.

A realistic simulation has high fidelity with respect to both the stimulus presented and the response elicited. It presents examinees with a stimulus that is very similar to the task stimulus that would actually be presented on the job, and it requires examinees to respond by performing a behavior very much as if they were on the job.

As simulations become less realistic, their task fidelity and response fidelity decrease. Simulations that have low task fidelity present examinees with only a written or oral description of the task stimulus, instead of a veridical representation. Simulations that have low response fidelity require examinees to give only a written or oral description of how they would handle the task presented, instead of actually requiring them to perform the response. Motowidlo, Dunnette, and Carter (1990) used the term "low-fidelity simulation" to refer to simulations that are low in both task and response fidelity.

The situational interview (Latham & Saari, 1984; Latham, Saari, Pursell, & Campion, 1980; Weekley & Gier, 1987) is a well known example of a low-fidelity simulation. Questions for the interview are developed based on a critical incident job analysis in which subject matter experts provide examples of behaviors that might be expected from outstanding, mediocre, and poor performers. Interviewers ask examinees to imagine they are in the hypothetical work situations and describe how they would respond to them. Validity estimates for this interview format have ranged from .14 (Latham & Saari, 1984) to .46 (Latham et al., 1980).

The situational inventory is another form of low fidelity simulation. It is similar in many respects to the situational interview. Instead of following an interview format with open-ended questions, however, the situational inventory is a paper-and-pencil instrument with multiple-choice response options that are objectively scored. An example of such an instrument is the Situational Judgement Test developed for Project A/Career Force (Campbell & Zook, 1992). Several tests of supervisory ability developed during the 1940s and 1950s, such as the How Supervise? test (File, 1945; File & Remmer, 1971; Rosen, 1961), the Leadership Evaluation and Development Scale (Mowry, 1957; Tenopyr, 1969), and the Supervisory Practices Test (Bruce & Learner, 1958), also include items that resemble some features of the situational inventory.

More recently, Motowidlo et al. (1990) developed and validated a low fidelity simulation in the form of a situational judgment test for the telecommunications industry. They developed item stems portraying hypothetical work situations from a critical incident job analysis of management positions. Management incumbents wrote descriptions of how they would handle the hypothetical work situations. Their descriptions were used to develop multiple-choice response alternatives for each situational item. Next, a sample of experienced managers evaluated the relative effectiveness of each response alternative and their judgements formed the basis for a scoring key. In a sample of approximately 120 management incumbents, Motowidlo et al. (1990) found that scores on their situational inventory correlated from .28 ($p<.01$) to .37 ($p<.01$) with supervisory ratings of job performance.

Motowidlo and Tippins (1993) reported two studies replicating these results. In one study, they reported an average predictive validity of .25 for the short version of the situational inventory. In the second study, they developed another situational inventory, this time for

marketing incumbents and reported an average concurrent validity estimate of .20 in a sample of 109 to 128 marketing incumbents from four telecommunications companies.

Combining results reported by Motowidlo et al. (1990) with those reported by Motowidlo and Tippins (1993) yields an overall validity estimate of .26, with race and sex differences estimated at less than a third of a standard deviation. These results confirm the potential usefulness of the low fidelity simulation in the form of situational inventory for employee selection. They show that simulations do not need to be high fidelity representations that require applicants to actually carry out some action to deal with a veridical task stimulus in order to be valid. Hypothetical task stimuli described in a few written sentences and hypothetical responses indicated by selection of multiple-choice alternatives can also predict job performance.

In summary, we decided to develop a written, multiple-choice test according to the situational inventory format. We planned to score the effectiveness of responses on this test and analyze relations between these effectiveness scores and other, presently available predictor constructs as well as job performance of NCO incumbents. This procedure was intended to maximize the efficiency with which we could assess tacit knowledge as well as constructs such as practical intelligence and social intelligence, for which no effective, separate measure currently exists. We decided to attempt to separately measure self-efficacy and priority setting skills, but within the same situational judgment framework. As our work progressed, we decided that it was not feasible to construct separate measures of the single constructs, but that it would be possible to more systematically specify the content of the situational inventory, thereby enhancing the understanding of such tests and enabling efficient construction of parallel forms. This systematic specification would focus on the identification of a taxonomy of problems derived from an analysis of the supervisory situations described by Army NCOs.

Development of the Army Leadership Questionnaire

This section reviews our development of a situational judgment test for NCOs, alluded to above, and referred to as the Army Leadership Questionnaire (ALQ). We used an iterative process to develop the ALQ. The development process is illustrated in Figure 2.1.

There were four steps in the development process:

1. Generate Problems. First, we developed an organizing framework to structure a problem writing exercise; 131 NCOs at two sites wrote approximately 1200 examples of situations that present a challenge to a leader.

2. Edit/Refine Problems and Define Problems Categories. Second, project staff read through all the problems, and, based on content, sorted them into 29 categories of NCO duties and responsibilities. We edited specific problems within the categories to form 195 problem stems that represented the categories. The 29 categories are listed and defined in Appendix A.

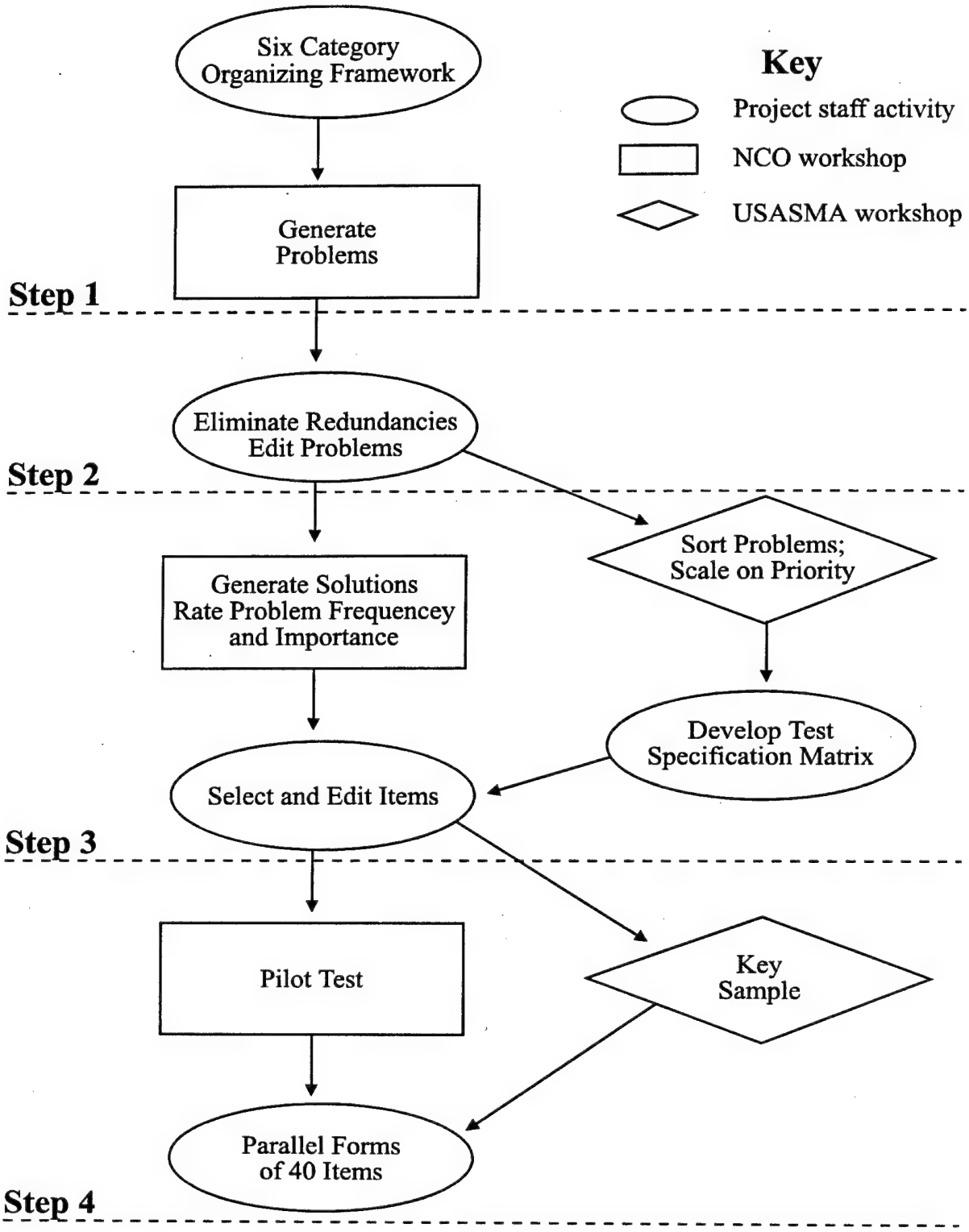


Figure 2.1
ALQ Development Process

3. Generate Problem Solutions and Write Items. NCOs (n = 280) at four sites wrote how they would respond to (subsets of) the 195 problem stems and also rated the importance and frequency of the problems. Then, with the aid of additional data from NCO students at the United States Army Sergeants Major Academy (USASMA), project staff developed a test specification matrix, and selected and edited 96 items (two forms of 48) for use in the pilot test.

4. Pilot Test, Key Parallel Test Forms, and Refine. Forms A and B of the ALQ were pilot tested with NCOs (n=186) at two different forts. NCOs at USASMA (n=135) also completed the instrument, with their answers providing a key for scoring the ALQ. Pilot test and keying sample data were used to refine the forms, producing two parallel forms of 40 items each.

A detailed description of the first three steps is provided below. Step four is reviewed in detail in Chapter 4.

Step 1: Generate Problems:

The first step in constructing the ALQ was to generate a pool of supervisory problems faced by NCOs. We scheduled NCO workshops at two sites to collect these data. NCOs in these workshops were to:

- think about a problem that required action on the part of an NCO
- write about the problem
- describe the problem
- describe what was done to address the problem
- rate the difficulty of the problem

These workshops followed a modified critical incident workshop format and agenda (Flanagan, 1954) to collect descriptions of problems. Specifically, rather than focussing these workshops on obtaining descriptions of the behaviors in the situations, we were primarily interested in generating rich descriptions of the problems themselves.

To prepare for the workshops, we developed a set of participant handouts, data collection materials, and a script to use in collecting supervisory problems. We also prepared advance copies of all materials and assembled them into a Research Instrument Review and Approval Summary (RIRAS) for ARI review.

At the workshops, participants listened to an overview of the project and their role in it, read the Privacy Act, and filled out a background information form. Then they were introduced to the notion of problems and responses and given examples and guidelines for writing problems. We asked them to respond to the following instructions:

Think about a time when you or somebody you know had to deal with a specific leadership problem.

1. *What exactly was the problem?*
2. *How did this person deal with it?*
3. *How difficult was the problem?*
   (rated on a 7 point scale: 1= easy; 7 = difficult)

To help NCOs in organizing their problem writing, we decided to provide them with a framework for development of NCO problems. Using our analysis of NCO tasks completed during the base period, we developed a parsimonious, six category framework (Table 2.1). These categories describe in general terms the kinds of actions that supervisors in levels E-5 through E-8 might take when faced with problems.

Table 2.1.
Framework for Development of NCO Problems

1.	Motivate subordinates
2.	Train subordinates
3.	Adapt to new and changing situations
4.	Counsel and discipline subordinates
5.	Make good decisions
6.	Plan and organize

Note: This category system was organized to represent reactions of NCOs to problems; the categories were derived by expert opinion from knowledge of typical reactions of supervisors across a wide variety of jobs. NCOs were therefore asked to look at each item on the list and think of problems that required these types of NCO actions.

We asked NCOs to write two examples from their own experiences for each category. Each NCO was encouraged to write problems individually. To ensure that NCOs understood the examples and were writing appropriate problems, we circulated around the room and reviewed their work while they each wrote their set of problems.

We requested participation of 72 NCOs in supervisory positions at Fort Benning and 90 NCOs at Fort Eustis, to yield approximately equivalent representation of Army MOS (basic groupings of "Combat", "Administrative/Support", and "Electronic/Repair"), and 2) ranks (E-5, E-6, and E-7/8). However, because Fort Benning had predominantly combat MOS available, we requested more non-combat personnel from Fort Eustis to compensate for this inequality.

We conducted six workshops over a three-day period (December 13 - 15, 1993) at Fort Benning, one day with each of the requested ranks (E-5, E-6, and a combination of E-7 and E-8). We also held six workshops in a similar fashion at Fort Eustis on January 10 - 12, 1994.

We collected a total of 1,204 problems from the 131 NCOs who participated in the situation generation workshops at Forts Benning and Eustis. Approximately half of this total was collected at each site. The rank and job type breakout for the sample of 131 NCO workshop attendees is shown in Table 2.2.

Step 2: Edit/Refine Problems and Define Problems

To assist us with refining and editing the problems collected in the workshops, we first placed the problems into categories. Three project staff members read and sorted approximately one third of the 1,204 problems on the basis of content similarity. Then we discussed the three staff members' different categorization schemes, identified the essentially equivalent categories, and resolved differences in how to treat the remaining specific kinds of problems and the categories they comprised. As a result of this sorting exercise, we formed a set of 29 "NCO

Table 2.2.
Problem Generation Workshops: Numbers of NCOs by Rank, Site, and MOS Grouping

Rank	Combat	Administrative/ Support	Electronic/ Repair	Total
E-5				
Fort Benning	20	3	0	23
Fort Eustis	2	12	3	18*
E-6				
Fort Benning	19	1	1	21
Fort Eustis	8	13	5	26
E-7/E-8				
Fort Benning	18	2	1	21
Fort Eustis	7	11	4	22
Total	74	42	14	131*

*Note: MOS grouping was missing for one E-5 at Fort Eustis.

problem Categories." We developed a working definition for each problem category, then all project staff reviewed them and suggested revisions. We divided up the total set of problems and, using the definitions, resorted the whole set such that every problem was uniquely assigned to only one category. The titles for the NCO Problem Categories are listed in Table 2.3.

Next, we needed to prepare the problems for presentation to additional NCOs who would be asked to write solutions to the problems. Specifically, we needed to eliminate redundant problems and edit the problem statements so that they were clear and complete. There were far too many problems to present to NCOs for writing problem responses. To develop a useful set of responses to a single problem for the test, we required several dozen individual responses written by NCOs. Therefore, we narrowed the set of problems through a series of sorting and

editing steps. First, we divided all the problem categories among staff members and wrote situational stems. Our procedure for doing this included the following steps:

- read through all problems in the category,
- identify the common or salient features of the problems within the category, and
- draft a problem stem that represents each of the salient features.

An expert reviewer reviewed all draft problem stems, looking for similarity in format, tone, level of detail, etc. We revised the draft problem stems and formed a final set of problem stems to represent the content of the total problem set. At the end of the problem editing process, we had a set of 195 draft problem stems. An example problem stem is:

You were just about finished conducting an accountability inspection of several pieces of equipment and found an item missing. The soldier claims that the item is not missing, then rolls his eyes backwards, turns his head away, and mumbles something as he walks away. You would...

Step 3: Generate Problem Solutions and Write Items

There were four major activities within this step. They included the following:

- Conduct Response Generation Workshops,
- Conduct Scaling and Taxonomy Development workshops,
- Develop a test specification matrix, and
- Select and refine items.

The purpose of the Response Generation Workshops was to obtain various potential and viable solutions to the 195 problems resulting from Step 2. NCOs of various ranks participated in the Response Generation Workshops. The purpose of the Scaling and Taxonomy Development Workshops was to order the problems along some meaningful constructs and to create a problem taxonomy based on these constructs. High ranking NCOs who were students at USASMA participated in the scaling and taxonomy development workshops. The scaling and taxonomy development workshops provided us with the information we needed to develop a test specification matrix. Then, we used the test specification matrix to select items for the ALQ. Finally, we used the information gathered from Step 2 and the Response Generation Workshops to create and edit the items. Each of these activities is described in detail below.

The goal of the Response Generation Workshops was to develop a pool of possible responses to the problems. We also wanted to obtain importance and frequency ratings on the problems and to verify the comprehensiveness of our categorization scheme. We conducted a series of workshops to collect these data. NCOs in these workshops were to:

- write responses to a subset of the 195 problems (problems were split into three sets so that each NCO responded to approximately 65 problems),

- rate the frequency and importance of each of the 29 categories, and
- rate the frequency and importance of individual problems (Note that only the E-7 and E-8 participants provided these ratings).

We prepared a set of participant handouts, data collection materials, and a script to use in collecting responses to supervisory problems. We also prepared advance copies of all materials and assembled them into a Research Instrument Review and Approval Summary (RIRAS) for ARI.

Conduct Response Generation Workshops

Included among these handouts were the response booklets. To prepare these, we randomly assigned the problem stems to three forms and distributed the problems such that no two adjacent problems were from the same problem category[1]. Then we created two versions of each form by reversing the problem order. Therefore, we ended up with six different response booklets: A1, A2, B1, B2, C1, and C2. Each booklet contained approximately 65 problems.

At each workshop, we described the project background, passed out the Privacy Act, and asked participants to fill out a background information form. We described leadership problems and responses, using examples. We passed out response booklets and asked NCOs to write brief explanations of how they would handle each of the problems described in their booklets.

We also asked senior-level NCOs (E-7s and E-8s) not only to write responses to problems, but also to provide judgments of frequency and importance for the problems. We used the following scales:

- "How often would an NCO be in a situation like this on the job?" (rated on a 6 point scale: 0 = An NCO would never face a situation like this on the job; 5 = Very Often)
- "Considering each problem as it stands, how important is it for an NCO to effectively handle a situation like this?" (rated on a 5-point scale: 1 = Not at all important; 5 = Extremely important)

To assist us with verifying the comprehensiveness of our problem categorization scheme, we asked all of the NCOs to review the problem categories and their definitions for completeness and accuracy. Then they rated frequency and importance at the category level: "How often would an NCO face situations like this on the job?" and "How important is it for an NCO to effectively handle situations like this?" We used this exercise to break up the writing task.

After the NCOs had completed their ratings and the solution generation task, we debriefed them. We asked them if there were any categories missing, or any types of problems that they view as important that they did not see in the solution generation task. We also asked them how useful items like these would be in predicting performance of NCOs. We recorded their suggestions and specific problem examples.

We conducted Response Generation Workshops at four sites. Two sessions were held per day for three days at each of the sites. We requested that POCs at each site provide us with NCOs distributed evenly across MOS (basic groupings of "Combat," "Administrative/ Support," and "Electronic/Repair" MOS), and rank (E-5, E-6, and E-7/8).

Table 2.3 .

Category Titles for NCO Problem Categories.

1. Emotional Health Problems
2. Legal Problems
3. Activity Planning
4. Extreme Weather Conditions
5. Equipment Problems (New, Broken, or Missing)
6. Poor Subordinate Appearance
7. Poor Billet Appearance
8. Inadequate Resources
9. Changes in Work Methods and Procedures
10. Conflict With Chain of Command
11. Subordinate Having Difficulty Adapting
12. Gender/Racial and Fairness Issues
13. Subordinate Unprepared for Mission
14. Poor NCO/Superior Performance
15. Poor Subordinate Conditioning
16. Subordinate Personal Problem(s)
17. Staffing Decisions and Problems
18. Supervisor-Peer Role Conflict
19. Subordinate Promotion/Award Decisions
20. Alcohol-Related Problems
21. Subordinate Financial Problems
22. Disrespect/Refusal to Follow Orders
23. Subordinate Poor Attitude or Behavior
24. Command Change
25. Poor Subordinate or Unit Effort
26. Extra Duty Requirements -- Difficult/Demanding Hours
27. Tardiness
28. Bureaucratic Snafu (a.k.a. System Failure)
29. Subordinate Career Uncertainty

The sites and the dates of the workshops were as follows:

Fort Drum	January 25 - 27, 1994
Fort Sill	February 22 - 24, 1994
Fort Knox	March 21 - 23, 1994

Fort Lewis April 18 - 20, 1994

A total of 300 NCOs participated. The division by rank and MOS grouping for this sample is shown in Table 2.4.

When all of the Response Generation Workshops were completed, project staff reviewed the NCOs' suggestions for new categories. We decided that no new categories were needed, because all the suggested categories and specific problems could be grouped with the existing NCO Problem Category scheme.

We also reviewed the other suggestions made by NCOs during their debriefing. One suggestion that came up several times was that an important aspect of NCO problem solving that is perhaps not captured in the items is the relative priority of the problem. NCOs told us that the manner in which they decide to deal with a problem is affected by its priority relative to other problems that they must deal with at that time. We decided to attempt to incorporate this suggestion into our development process.

At a February 1994 progress meeting with ARI staff, we reviewed and discussed the results of the Response Generation Workshops. The decision was made to proceed with development of a paper-and-pencil instrument designed to assess NCO tacit knowledge using a situational judgment format. The format for each item would include a 2-3 sentence

description of a problem, and a listing of 4-5 solution options. We called this instrument the Army Leadership Questionnaire (ALQ).

Conduct Scaling and Taxonomy Development Workshops

To continue with development of the ALQ, we needed additional data. Previous workshops had provided us with the content for the instrument, but we needed to have some structure for choosing items. A structure for selecting and creating items is important, because it improves the representativeness and comprehensiveness of the items according to important variables. We decided to structure the ALQ according to a test specification matrix in which the rows would be defined by a qualitative categorization of the problem domain (i.e., type of problem), and the columns by an ordering of the problems in terms of urgency or "priority." Problem priority would be useful in selecting items because opinions expressed by NCOs in our workshops strongly indicated that perceptions of priority affect NCO problem solving behavior. The test specification matrix would be more sound if the values on these dimensions were determined from data collected from SMEs. Therefore we conducted a series of workshops to collect these data from senior NCOs attending USASMA at Fort Bliss.

This series of workshops was conducted to scale the problems and to place them into a useful taxonomy. We conducted three workshops in this series. One of the workshops, "Problem Sorting," was conducted to obtain data for developing a parsimonious set of empirically-based situational categories. The other two workshops, "Multiple Rank Ordering" and "Rating Frequency and Importance," allowed us to develop a unidimensional priority score

for each problem that we would use in developing our assessments. These workshops are discussed below.

To prepare for these workshops, we needed to choose the problems that would be used. Although the purpose of the workshops was to scale and categorize all of the problems presented in the response generation phase, we recognized that the workshop tasks we were considering would be unworkable if all of the problems were included. So, we decided to select a subset of problems that represented the domain of problems faced by NCOs.

Table 2.4.

Response Generation Workshops: Numbers of NCOs by Rank, Site, and MOS Grouping

Rank	Combat	Administrative /Support	Electronic /Repair	Total
E-5	57	38	36	**131**
Fort Drum	10	12	11	
Fort Sill	20	3	9	
Fort Knox	17	15	15	
Fort Lewis	10	8	1	
E-6	28	25	13	**66**
Fort Drum	11	6	4	
Fort Sill	9	8	4	
Fort Knox	2	7	1	
Fort Lewis	6	4	4	
E-7/E-8	39	36	9	**83**
Fort Drum	10	9	2	
Fort Sill	9	13	3	
Fort Knox	9	10	3	
Fort Lewis	11	4	1	
Total	**124**	**98**	**58**	**280***

*Note: The following data were missing: Fort Drum - missing rank for 3 soldiers; Fort Sill - missing rank for 7 soldiers, 1 soldier was E-4, and 1 E-5 was missing MOS; Fort Knox - missing rank for 4 soldiers, 1 E-6 was missing MOS; Fort Lewis - missing rank for 1 soldier, 2 soldiers were E-4s.

To ensure that we were choosing problems that represented the domain of potential problems, we based our selection on the NCO Problem Category scheme discussed earlier. To make it easier to work with this scheme, we collapsed the 29 categories down to 12 higher order

NCO Problem Categories. We then selected problems from each of these higher-order categories in proportion to the number of problems in that category. We selected problems that:

- were plausible,
- dealt with a variety of issues,
- varied in apparent urgency, and
- appeared to have a non-obvious solution.

We selected a total of 65 problems for these workshops. All 65 problems were used in both the Problem Sorting Workshop and the Rating Frequency and Importance Workshop. A subset of these problems (k=41) was used in the Multiple Rank Ordering Workshop.

A total of 97 senior NCOs participated in the workshops. Table 2.5 shows numbers of participants by rank and MOS grouping.

Table 2.5.

NCO Academy Workshops: Numbers of NCOs by Rank and MOS Grouping

Rank	Combat	Administrative/ Support	Electronics/ Repair	Total
E-8	35	43	14	92
E-9	4	1	0	5
Total	39	44	14	97

The data collection activities for each workshop are listed below.

Problem Sorting Workshop

The purpose of this workshop was to provide us with data for developing leadership problem categories. There are many ways of determining a category structure, but most of these methods require more SME time than we had available. Therefore, we decided to use a sorting and data analysis technique refined by Rosenberg and Kim (1975). Using this technique, subjects receive a set of cards with the stimuli (i.e. problems) printed on them. They are instructed to sort the stimuli into piles, where stimuli within a pile are similar to one another. Subjects are allowed to use any number of piles.

Fifty sets of the 65 problems selected for this workshop were then printed onto 3 x 5 cards. We also prepared written and oral instructions. Subjects (n=48) were asked to sort the problems into piles similar in priority. All subjects completed the sorting activity.

Data from the Problem Sorting Workshop were converted to a 65 x 65 similarity matrix. Formulas for developing the similarity matrix are presented in Rosenberg and Kim (1975). This similarity matrix was then factor analyzed and a varimax orthogonal rotation was applied to six of the factors. This six factor summary of the 65 problems served as a basis for categorizing the remaining 130 problems. The 130 problems were sorted into the six categories, plus two theoretically developed categories for problems not fitting one of the six. The eight categories and their definitions are listed in Table 2.6. These eight categories were later used to assist us in selecting problems for the ALQ and the Leadership Problems Inventory. There were also useful in defining additional criterion.

Table 2.6

ECQUIP Situational Categories from Factor Analysis of Simple Similarity Coefficients Based on Card Sort Data Gathered at the NCO Academy

1. Routine Requests or Complaints: Subordinates complain about or make routine requests for things such as a different assignment, a transfer, time off, or advice.
2. Conduct Problems: Subordinate are in trouble because of problems such as violence, drug or alcohol abuse, or financial mismanagement.
3. Shortage of Resources: Unit resources such as equipment, time, or training opportunities are short or unavailable.
4. Neglect of Responsibilities: Subordinates habitually neglect responsibilities in ways such as being late for formations, taking excessive sick calls, failing to learn basic skills, or not maintaining good physical condition or neat barracks conditions.
5. Personal Problems: Subordinates have personal problems through no fault of their own, such as the death of a friend or relative, or errors or inefficiencies caused by Army.
6. Complaints about Command Actions: Soldiers complain or react negatively to action taken by people in the chain of command.
7. Skill Deficits: Subordinate lack necessary skills, knowledge, or experience.
8. Subordinate Recognition or Promotion: Subordinates are candidates for promotion or recognition for exemplary performance.

Note: Categories 7 and 8 were not derived through the factor analysis. They were developed to allow us to categorize those problems that did not fit into one of the six empirically derived categories.

Multiple Rank Ordering Workshop. The purpose of this workshop was to develop a unidimensional priority score for each problem. Data from this series of workshops were used in three different ways to derive priority scores for all 195 problems. The data collected in this particular workshop were used in each of the methods. Figure 2.2 illustrates these methods. For those problems included in the Multiple Rank Ordering Workshop (k=41), we calculated priority scores directly using the method of multiple rank ordering (Gulliksen & Tucker, 1961), described

below. Then we regressed these scores onto the frequency and importance scores obtained in other workshops. For the remainder of the problems, we predicted the priority scores based on these equations. For example, for those problems that were included in the USASMA frequency and importance workshop (k=24), we predicted priority scores from frequency and importance ratings obtained at the Response Generation Workshops and the USASMA Frequency and Importance Ratings Workshop. For the other 130 of the problems, priority scores were predicted using the frequency and importance data gathered at the Response Generation Workshops only.

The multiple-rank order (MRO) method developed by Gulliksen and Tucker (1961) is a means of obtaining paired-comparison data with a reduction in subject labor. With the MRO method, subjects rank-order stimuli within subsets or blocks. The subsets are constructed such that paired-comparison data can be obtained for all possible pairs of stimuli using the within-block rank-order data. Interval scale values are then obtained through the law of comparative judgment (Guilford, 1954).

We chose only 41 problems for this workshop because of the limited number of available SMEs and the short time available for the scaling task. The 41 problems were combined in blocks of five problems each to create unique problem blocks. The 41 problems required 820 paired comparisons; thus, each problem required 10 paired comparisons, and we created 82 unique problem blocks. We created 50 "MRO exercises" which were completely balanced, meaning that they were designed so that every problem was paired once with every other problem, and no two problems occurred together more than once. For each of these 50 MRO exercises, we randomized the order of blocks within the exercise and the order of problems within the blocks. Each of these MRO exercises was then divided into two booklets (41 problem blocks in each booklet). Each subject received a booklet of 41 problem blocks. Subjects were instructed to rank the problems in each block in terms of priority. As subjects completed a first booklet, they were given an additional booklet to complete. Several subjects completed only one booklet, while those who worked quickly completed a second or even a third booklet until the fifty MRO exercises (100 booklets) were completed. We converted the subjects' responses to pair-wise comparison data.

We applied standard scaling procedures (Torgerson, 1958), to produce a unidimensional "priority" scale score for each of the problems included in the exercise. First, a 41 x 41 matrix of preferences was formed by placing the problem identification numbers along the columns and rows. The value in each cell indicates the proportion of times the problem identified on the column heading was ranked as a higher priority than the problem identified for that row. This matrix was then converted by taking the inverse of the cumulative normal distribution for each point in the first matrix. The mean of each column signified a priority score. For ease of use, these scores were converted to "T scores" (M = 50, S.D. = 10). Reliability analyses indicated that paired-comparison results were consistent across respondents, and the priority scores from this process were highly reliable (KR-20 r=.98).

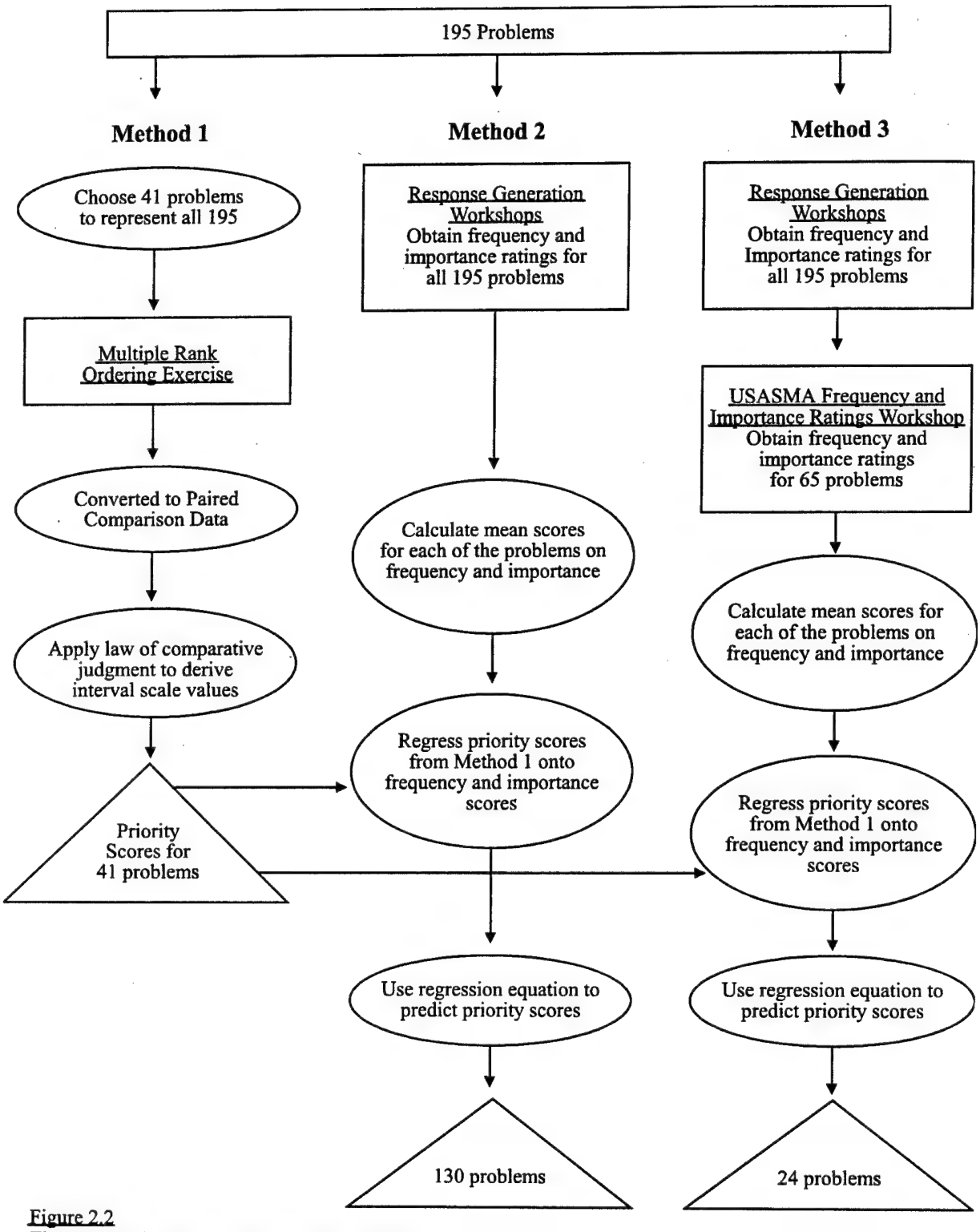


Figure 2.2
Three Methods Used to Derive Priority Scores

Frequency and Importance Ratings Workshop. Participants who completed either the Multiple Rank Ordering Workshop or Problem Sorting Workshop were then asked to complete the Frequency and Importance Rating Workshop. The purpose of this workshop was to provide additional data for developing a unidimensional scale of problem priority. Participants received a questionnaire with problems selected from those used in the Response Generation Workshops. On the questionnaire, participants responded to the following questions:

How often would an NCO be in a situation like this on the job?

Considering each problem as it stands, how important is it for an NCO to effectively handle a situation like this?

The first question was to be rated on a 6 point scale (0 = An NCO would never face a situation like this on the job; 5 = Very Often), and the second question was to be rated on a 5-point scale (1 = Not at all important; 5 = Extremely important).

The questionnaire included the same 65 problems used in the Problem Sorting Workshop. A total of 86 NCOs completed the ratings. The KR-20 reliability coefficients for the mean of their frequency and importance ratings were .97 and .97 respectively, indicating that problems were rated consistently across subjects.

Frequency and importance scores for each of the problems used in the Response Generation Workshops were created by taking the mean of each problem's frequency and importance ratings. For most of the problems (i.e., 130 of them), these scores were based on data collected at the response generation sites only. For the remainder of the problems, these scores were based on data collected at the NCO Academy and the response generation sites.

As indicated earlier, we then used these scores to predict the priority scores for the problems not included in the Multiple Rank Ordering Workshop. We regressed the priority scores from the MRO workshop onto the frequency and importance scores calculated above. Two regressions were calculated. In the first case, priority scores were regressed onto frequency and importance data gathered at the response generation sites only. In the second case, priority scores were regressed onto frequency and importance data gathered at the NCO Academy and the response generation sites. We used the parameter estimates from these regressions to develop a predicted priority score for all of the problems not used in the Multiple Rank Ordering Workshops. Two regression equations were needed because we did not have frequency and importance data from both the NCO Academy and the Response Generation Workshops for all problems, and the prediction equations differed depending on the source of the data.

The regression analyses are provided in Table 2.7. The parameter estimates for the frequency variable were negative for both regressions. This finding was counter-intuitive, especially given that the parameter estimates for the importance variable were positive, as would be expected. We attribute this result to the high intercorrelations of frequency and importance ratings. For both sets of regressions, frequency and importance accounted for over 60 percent of the variance in the priority score. Therefore, parameter estimates from these regressions could be

used to obtain reasonable estimates of the priority scores for all 195 problems. Parameter estimates from the first regression were used to derive predicted priority scores for 24 of the problems. Parameter estimates from the second regression were used to derive predicted priority scores for the remaining 130 problems.

Table 2.7
Predicting Relative Priority Using Frequency and Importance Ratings.

Multiple Regression Results

Source	df	SS	F	R^2	Adj-R^2
Model	2	1615.30	75.26	.79*	.79*
Error	38	815.54			
Total	40	4046.15			

Parameter Estimates

Variable	df	Parameter Estimate	Standard Error	T for H0: Parameter = 0
Intercept	1	-9.53	7.37	-1.29
Frequency	1	-4.25	1.59	-2.65*
Importance	1	18.21	1.58	11.50*

* $p < .05$

Data Source: NCG Academy plus Response Generation Workshops

Multiple Regression Results

Source	df	SS	F	R^2	Adj-R^2
Model	2	2495.80	30.58	.62*	.59*
Error	38	1550.35			
Total	40	4046.15			

Parameter Estimates

Variable	df	Parameter Estimate	Standard Error	T for H0: Parameter = 0
Intercept	1	23.873	11.14	2.14*
Frequency	1	-15.386	2.40	-6.39*
Importance	1	16.893	2.91	5.81*

* $p < .05$

Data Source: Response Generation Workshops

Develop a Test Specification Matrix

To facilitate selection of problems for the ALQ, we developed a two-dimensional test specification matrix. This matrix was developed using the data gathered in the scaling and taxonomy development workshops. The matrix consisted of the eight categories (from the Problem Sorting Workshop) listed along the rows, and three groups of problems arranged by priority score (High, Medium, and Low) listed along the columns. The three groups were

derived by trichotomizing the problems according to priority score. Each problem was placed in one of the matrix cells. This 8 x 3 test specification matrix was also used to develop another predictor instrument that we will discuss later: the Leadership Problems Inventory.

Select and Refine Items

When the Response Generation Workshops were completed, we put all the responses for each problem together and grouped the problems according to the NCO Problem Category (1 through 29). At that point, we had multiple responses (approximately 80) for each of 195 problems.

We then made item writing assignments. Each project staff member was given one or more problem categories and asked to write items for all of the problems within that category. No changes were made to the problem stems at this point, as they had already been edited prior to selecting the responses (already written by NCOs) that we would use as response options in ;the item. The process included the following steps:

- Review the problem and ensure that you understand it.
- Read all of the responses one time through.
- Sort the responses into piles, placing similar responses together.
- Characterize the responses within each pile in a written statement. This statement should be:
    - action oriented, and thus begin with a verb
    - short--typically only one sentence. Responses should never be more than three sentences
    - of equal length for a given problem
    - of equal scope for a given problem—a response with several actions included in it would be favored too often, and thus would decrease the variance of responses on the item
    - describe an action, or set of actions that is independent of those found in other statements
    - written so that the rich contextual nature of the response is retained
    - plausible to the examinee as a response option
- Select five response statements that represent the broadest variety of responses possible. Also, try to satisfy as many conditions as possible from the above list for each of these response statements.

Figure 2.3 presents an example item. Note that examinees are to provide two responses for each item. They are to indicate the "most effective" solution, and the "least effective" solution. This dual response format was chosen to maximize the amount of information that is obtained for each problem. It was also thought that the domain of problem understanding includes knowledge of the least effective solution as well as the most effective solution.

Figure 2.3:
Example ALQ Item

Item

During a training course with your unit you notice that some of your soldiers actively participate, while others don't say a word. You would...

a. talk to those who do not participate and encourage them to speak out.

b. change the training plan so that participation is required.

c. make those individuals who don't talk teach the course.

d. let those who participate leave early and keep the others for more training.

e. do nothing; if they don't learn they will only be hurting themselves.

Response:

Most effective..... A B C D E

Least effective..... A B C D E

Once the first draft of each item was created, we were ready to select items for the ALQ. Two forms (48 items each) were created. Items for these forms were selected based on the 8 x 3 test specification matrix. Project staff reviewed items in each cell of the matrix and formed pairs of items that:

- had similar content,
- dealt with problems that were realistic,
- had five plausible alternatives, and
- were of reasonable length.

One item in each pair was then assigned to Form A, while the other was assigned to Form B. The number of pairs selected from each cell ranged from one to three and depended on the total number of items in that cell. The problem identification codes selected for each form are shown in Table 2.8.

Table 2.8

Problems Selected for the ALQ

	Level of Priority					
	High		Medium		Low	
Situational Category	Form A	Form B	Form A	Form B	Form A	Form B
Routine Requests or Complaints	C42	A28	C50	A57	C49	A10
	B64	B26	A2	C59	C35	B51
Conduct Problems	A44	C6	C32	C54	B31	A38
	A24	B28	B35	A20	B3	C55
Shortage of Resources	A27	B66	A49	B48	C11	B44
	C64	B46	B62	B10	A8	A13
	B61	C40	A22	A7	B65	B58
Neglect of Responsibilities	C28	B2	A64	C47	B33	B8
	C34	C59	B42	C1	C14	B16
Personal Problems	A1	C3	B36	B6	A26	A59
	B53	C44	A32	A14	C58	B22
Complaints About Command Actions	A9	C61	B47	B15	A53	A3
	B50	A63	B45	C38	B55	C20
Skill Deficits	B30	A62	A21	C22	C48	C63
Subordinate Recognition or Promotion	C15	A12	B37	B43	A4	B11

Note: Each cell contains the identification numbers for the problems selected for that situational category, level of priority, and instrument form (A or B). Identification numbers indicate the form letter (A, B, C) and problem number used on the Response Generation Forms. Problem identification numbers that appear in like positions for Forms A and B are paired with one another. Members of each pair were matched on the basis of judgments of content similarity.

Staff members then reviewed and edited the selected items again, ensuring that the response options were all clear and plausible. Then we wrote a set of instructions with an objective of keeping them short and simple. This set of instructions included an example of exactly how to respond to an item. We "tried-out" these instructions and the response format on several NCOs,

made some minor revisions, and then concluded that the measure was ready for pilot testing. The major steps in the development process included:

Development of a Measure of Self-Efficacy

The development of a measure of self-efficacy took a fairly straightforward path, as is illustrated in Figure 2.4.

The major steps in the development process included:

1. Development of a measurement strategy,
2. Pilot test, and
3. Addition of two new instruments that tap self-efficacy in different ways.

Only the first step is discussed in this chapter. The pilot test, and the subsequent revisions/additions to our self-efficacy measurement plans are discussed in detail in Chapters 4 and 5 respectively.

Self-efficacy is "...the belief in one's own ability to mobilize the motivational and cognitive resources and the behaviors that are needed to meet task demands" (Bandura, 1982). This construct is typically measured by asking respondents about their degree of confidence in completing a specific task (Lee & Bobko, 1994). We decided to measure self-efficacy through similar means, that is, by asking respondents about their degree of confidence in successfully completing a specific task.

However, there are two problems associated with using traditional methods of measuring self-efficacy in a selection setting. First, most applications of the traditional method have referenced only one, or very few tasks. For example, a recent study (Lee & Bobko, 1994) asked student subjects about their degree of confidence in achieving various grades (i.e., A, B, C, and D) in a particular class. We were concerned that self-efficacy measured in reference to only one task may not generalize to self-efficacy on other tasks. Our second concern was that the self-report nature of the traditional method would lead to a positive bias in the responses, and therefore a reduction of the variance obtained on the measure.

Therefore, we developed an approach to measuring self-efficacy that addresses both problems. Essentially, we planned an approach where examinees would indicate their degree of confidence in their responses to ALQ items. For each problem presented in the ALQ, the respondent would indicate (on a 10-point graphic scale) his/her degree of confidence in selecting the "most effective solution." Respondents would answer a similar question for the "least effective solution." An example of how this type of question is applied to an ALQ item is provided in Figure 2.5.

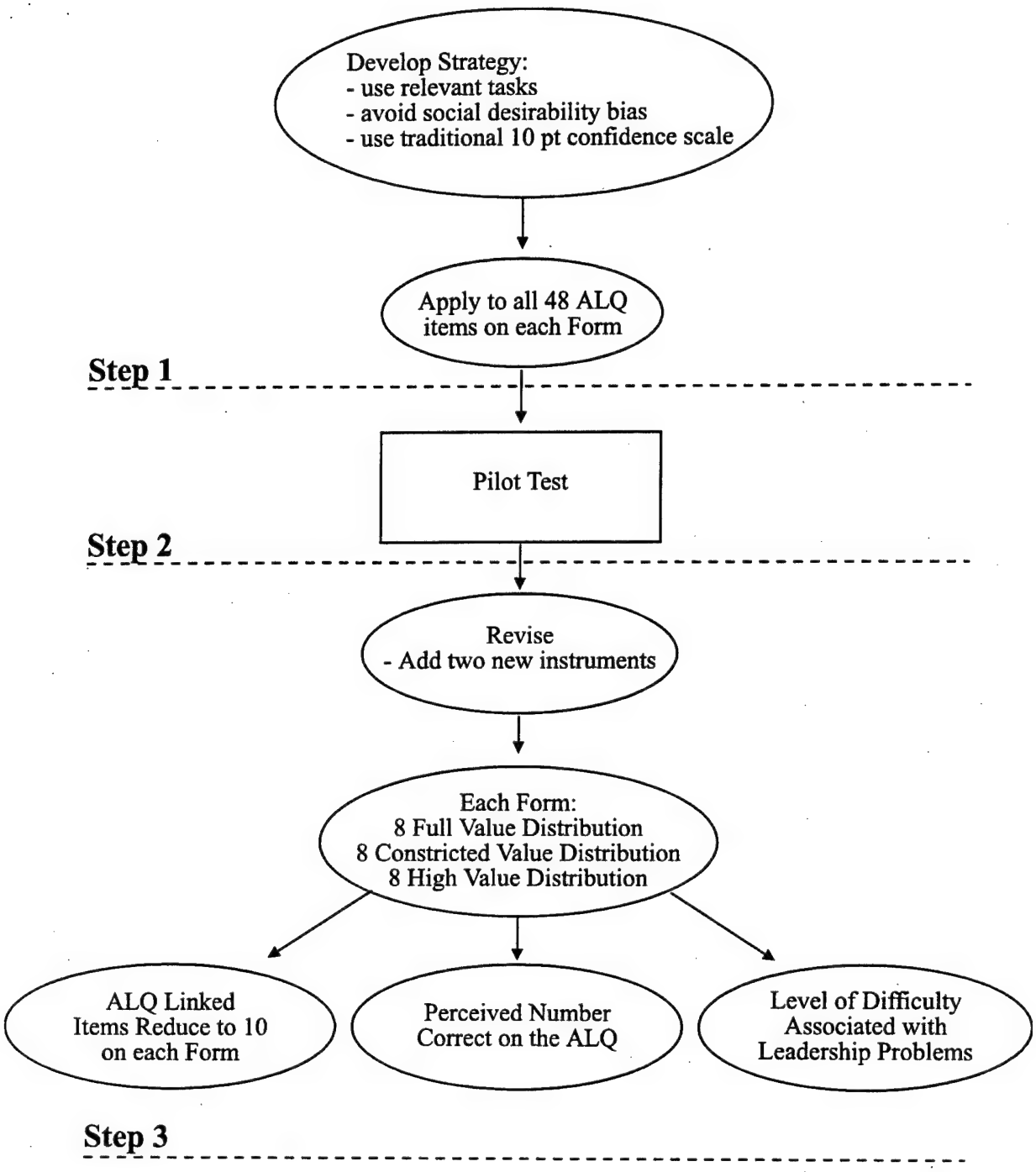


Figure 2.4
Self-efficacy measure development process.

Figure 2.5
Example of how self-efficacy scales are applied to ALQ items.

ALQ Item:

During a training course with your unit you notice that some of your soldiers actively participate, while others don't say word. You would...

a. talk to those who do not participate and encourage them to speak out.

b. change the training plan so that participation is required.

c. make those individuals who don't talk teach the course.

d. let those who participate leave early and keep the others for more training.

e. do nothing; if they don't learn they will only be hurting themselves.

Response:

	Choose the most effective and least effective respones.		**Rate your confidence that you picked the most effective and least effective responses.**
Most effective	..... A B C D E	.....	1 2 3 4 5 6 7 8 9 10
Least effective	..... A B C D E	.....	1 2 3 4 5 6 7 8 9 10

Through this method, we could obtain a measure of self-efficacy as it applies to two aspects of every problem in the instrument. Thus, as the ALQ was designed to cover the domain of problems faced by NCOs, this measure of self-efficacy would be more comprehensive than traditional measures of self-efficacy that present a restricted set of "tasks." In addition, the large number of items (96) would likely mean that the instrument would be highly reliable. Obtaining a rating on this many tasks would typically be too much of a burden, because the examinee would have to read a large number of tasks to provide ratings. In this case however, the examinee would have already read the item in order to respond to the tacit knowledge portion of the ALQ. Thus, the added burden to the examinee for obtaining the information would be relatively small.

There is another advantage to using ALQ items as tasks to measure self-efficacy. Positive bias could be reduced, as respondents might see it to their advantage to respond truthfully and without bias to the confidence ratings. This is because their actual performance on

the item could be compared to their confidence rating. Due to the various ways in which the confidence rating can be used, it is not obvious to the respondent what the "most favorable" self report rating would be. For example, respondents may have hesitated in assigning a high confidence rating to an item that they found difficult if they believed that items with a high confidence rating would be weighted more heavily.

After we settled on our measurement technique, we prepared the self-efficacy measure for pilot testing. We wrote a set of instructions-- again with the objective of keeping them short and simple. These instructions included an example of exactly how to complete a response on this measure. We "tried-out" these instructions and the response format on several individuals, made some minor revisions, and then concluded that the measure was ready for pilot testing.
a rating on this many tasks would typically be too much of a burden, because the examinee would have to read a large number of tasks to provide ratings. In this case however, the examinee would have already read the item in order to respond to the tacit knowledge portion of the ALQ. Thus, the added burden to the examinee for obtaining the information would be relatively small.

There is another advantage to using ALQ items as tasks to measure self-efficacy. Positive bias could be reduced, as respondents might see it to their advantage to respond truthfully and without bias to the confidence ratings. This is because their actual performance on the item could be compared to their confidence rating. Due to the various ways in which the confidence rating can be used, it is not obvious to the respondent what the "most favorable" self report rating would be. For example, respondents may have hesitated in assigning a high confidence rating to an item that they found difficult if they believed that items with a high confidence rating would be weighted more heavily.

After we settled on our measurement technique, we prepared the self-efficacy measure for pilot testing. We wrote a set of instructions-- again with the objective of keeping them short and simple. These instructions included an example of exactly how to complete a response on this measure. We "tried-out" these instructions and the response format on several individuals, made some minor revisions, and then concluded that the measure was ready for pilot testing.

Development of the Leadership Problems Inventory (LPI)

As our work progressed in the development of the ALQ, we continued to keep an eye out for additional ways of measuring the constructs recommended in the base report (Peterson et al., 1993). During the Response Generation Workshops, we asked NCOs to comment on our planned situational judgment test. It was through their feedback that we got the idea to develop the LPI.

As mentioned earlier, one comment came up several times during the response generation workshops; that is, the relative priority of a problem is an important aspect of problem solving that must be considered. Specifically, NCOs told us that the manner in which they decide to deal with a problem is affected by its priority relative to other problems that they must deal with at

that time. The base report (Peterson et al., 1993) supported this notion several skills that underlie the act of setting priorities were identified as important constructs to measure, including planning, problem understanding, and problem solving.

At the February 1994 progress meeting with ARI staff, we discussed the comments of the NCOs on this issue, pointed out the base report recommendations, and made an additional recommendation to develop a measure designed to test priority-setting. The decision was made to proceed with development of a paper-and-pencil instrument assumed to assess NCO planning, problem understanding, and problem solving skills. These skills would be assessed as they relate to the act of setting priorities. The format of the items would be simple. A respondent would view a set of two to seven problems (a "problem block") and determine which problem he or she would act on first, second, third, and so on. The instrument would consist of a series of these "problem blocks."

Our process for development of the LPI was interwoven with that for the ALQ. Many of the products created for the purpose of the ALQ also served the LPI. Therefore, we have already discussed much of the work needed to develop the LPI earlier in this chapter under the description of the ALQ development process. Here, we review Step 3 in some detail, as this step differed somewhat for the LPI. Step 4 is reviewed in detail in Chapter 4. The LPI development process is illustrated in Figure 2.6.

There were four steps in the development process:

Step 1: Generate Problems

First, we developed an organizing framework to structure a problem writing exercise; 131 NCOs at two sites wrote 1,204 examples of situations that present a challenge to a leader. See the discussion of this step under ALQ development. Products derived for the ALQ in this step were also used for LPI development.

Step 2: Edit/Refine Problems and Define Problems Categories

Second, project staff read through all the problems, and, based on content, sorted them into 29 categories of NCO duties and responsibilities. We edited specific problems within the categories to form 195 problem stems that represented the categories. See the discussion of this step under ALQ development. Products derived for the ALQ in this step were also used for LPI development.

Step 3: Obtain Priority Scores for Each Problem and Place Them Into Problem Blocks

There were three major activities within this step. They included:

- Conduct scaling and taxonomy development workshops
- Develop a test specification matrix, and
- Place problems into problem blocks

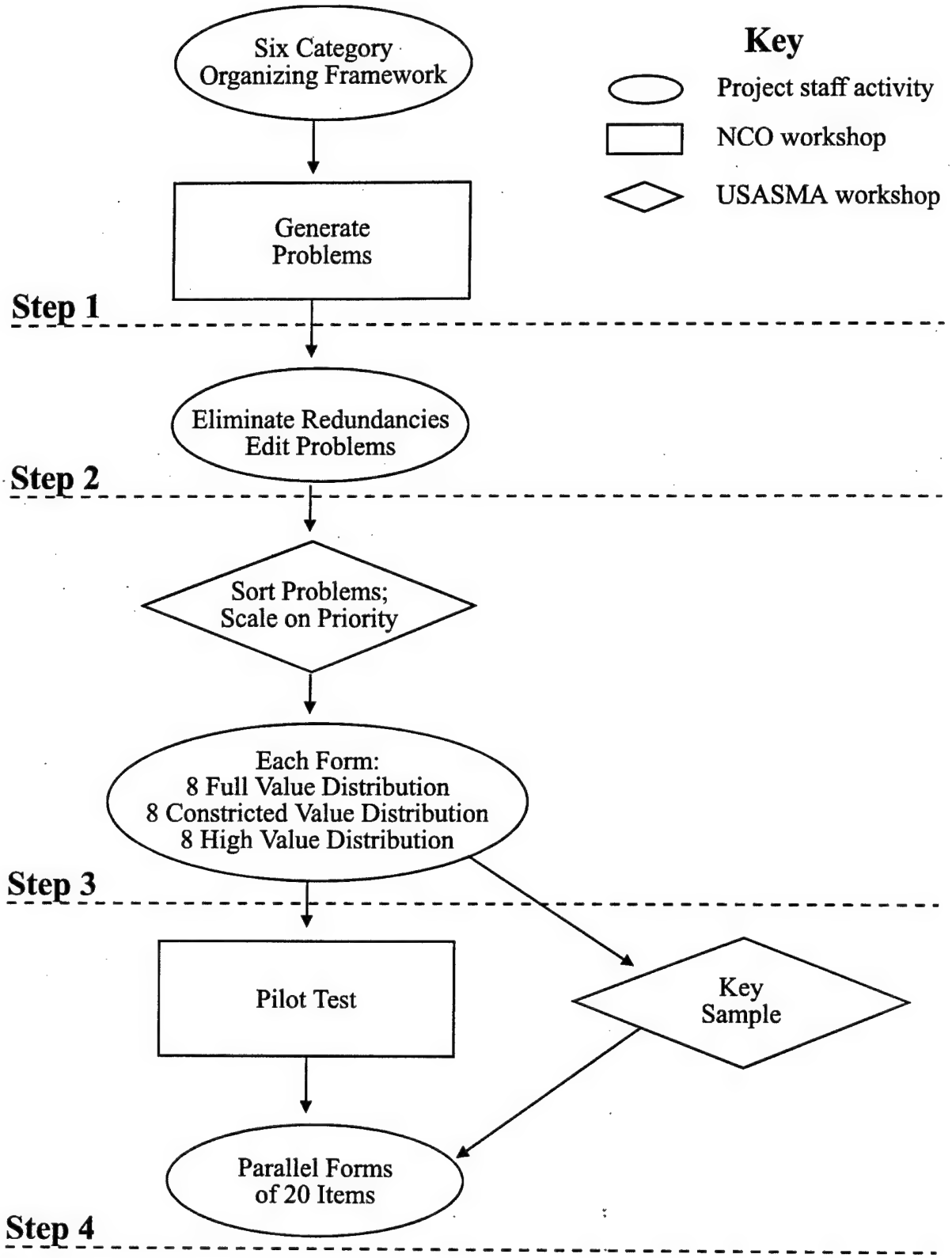


Figure 2.6
LPI Development Process

The first two activities were covered in some detail under our discussion of the development of the ALQ. We do not repeat that description here. We do, however, highlight aspects of these activities that led to the development of the LPI. The third activity is unique to development of the LPI, and thus we cover it in detail.

Recall that the purpose of the Scaling and Taxonomy Development Workshops was to scale the problems according to priority and to create a problem taxonomy. High ranking NCOs who were students at USASMA participated in these workshops. We then developed a test specification matrix from these workshop data. Scaled priority scores and the test specification matrix were then used to place problems into problem blocks. These problem blocks formed the actual LPI items. Conduct Scaling and Taxonomy Development Workshops

To continue with development of LPI, we needed additional data. Previous workshops had provided us with the content for the instrument, but we needed some structure for establishing problem sets. More importantly, however, we needed to obtain a priority score for each problem so that we could create problem sets for the LPI of varying levels of difficulty. It seemed reasonable to hypothesize that problem sets with homogeneous levels of problem priority would be more difficult than problem sets with heterogeneous levels of problem priority. Therefore, similar to the ALQ, we decided to structure the LPI according to the dimensions of priority score, and a qualitative categorization of the problem domain (i.e., type of problem).

Recall from our discussion of these workshops earlier in this chapter that there were three workshops:

- Problem Sorting Workshop
- Multiple Rank Ordering Workshop
- Frequency and Importance Rating Workshop

The Problem Sorting Workshop was conducted for the purpose of developing an empirically based category scheme for the problems. The other two workshops were used to scale each of the problems according to priority.

Recall that the response format in the Multiple Rank Ordering Workshop had the participants responding to blocks of problems, and assigning a priority to each of the problems within that block. This response format was, by no coincidence, identical to that eventually used for the pilot and validation versions of the LPI. We wanted to scale the priority of each problem in a context similar to the one we expected to use on the LPI. Context is important because the "priority of a problem" has no meaning unless it is considered relative to other problems. Therefore, we scaled priority using the method of multiple-rank order (MRO). As mentioned before, this method has several other positive features that merit its use in this setting.

Development of a Test Specification Matrix

The test specification matrix created for the development of the ALQ was also useful for the development of the LPI. However, the matrix was used differently here. Rather than pulling

problems strictly from the matrix according to both situational category and priority score, we focussed on choosing problems with particular priority scores. We also tried to choose problems from a wide variety of situational categories.

Placement of Problems into Problem Blocks

The purpose of this activity was to create the items for the LPI by choosing problems that would be placed together in problem blocks. Each item on the LPI consists of a set (or block) of 5 problems, and the respondent is asked to rank the problems in terms of priority. An example LPI item is provided in Figure 2.7.

We chose problems for the LPI to create three kinds of priority score distributions for each problem set. We assumed that problem blocks with homogeneous levels of problem priority would be more difficult than problem blocks with heterogeneous levels of problem priority. The three distributions are:

- High value distribution, where problems are only selected for these blocks if they have a priority score above the median.

- Constricted value distribution, where problems are only selected for these blocks if they have a priority score within one standard deviation of the mean.
- Full value distribution, where problems are selected to represent the full range of priority scores.

Items using blocks with a high value distribution are meant to characterize the most difficult priority setting tasks, where all of the problems have a high priority. Items using blocks with a constricted value distribution should also be difficult because all of the problems are similar in priority. Finally, we expect that items using blocks with a full value distribution would be easier, as there is inherently more difference between the problems in terms of priority.

To select problems for the LPI, we used the 8 x 3 Test Specification Matrix discussed earlier. Problems in each cell of the 8 x 3 test specification matrix were reviewed by project staff who selected pairs of problems that:

- fit with the distribution needs of the given problem set,
- had similar priority scores,
- had a direct, unidimensionally-scaled priority score (derived from the 41 pairs of items scaled in the MRO workshop) for at least one of the problems, and
- dealt with problems that, as a set, could be reasonably expected to co-occur.

Two forms of the LPI (24 items each) were created. One problem in each pair was then assigned to a set in Form A, while the other was assigned to a corresponding set in Form B. These problems were assigned so that at least 50 percent of the problems on each form had direct, unidimensionally-scaled priority scores.

Figure 2.7
Example LPI Item

First, please read all 5 problems.

A. One of your sergeants is a crew chief who has the reputation of keeping his mission-critical equipment clean and well maintained. This morning he showed up late and obviously intoxicated. One of his co-workers asked if he had been drinking and said he would get a replacement if he had. The sergeant said he had not been drinking and was fine. The co-worker comes to you and repeats the conversation he had with your sergeant.

B. The wife of a soldier in your platoon calls and says her husband is causing a disturbance at home and she needs help, but she does not want to involve the Military Police (MPs) unless necessary.

C. A soldier in your platoon reports for work with alcohol on his breath. His is very loud. Other soldiers tell you he came to work drunk.

D. The non-commissioned officer (NCO) who was in charge of a firing range became ill and you were assigned at the last minute to fill in for him. You have never done this before and you only have one day to prepare for the range.

E. A new private has finished basic training. He reports to your unit and tells you that he has a big problem. He has a wife and kids and does not have a place to live, and does not have enough money for rent or food.

Next, indicate the order in which you would handle each of the problems. In the "1st" box write the letter corresponding to the problem highest in priority; write the letter for your second ranked problem into the "2nd" box, and so on.

PRIORITY

1st	2nd	3rd	4th	5th

Highest Priority — Lowest Priority

(Use all five letters, "A" through "E," and be sure to write each letter in only one box.)

The number of problem blocks for each priority distribution are shown in Table 2.9. A summary of the distribution of these problems is shown in Tables 2.10 and 2.11. Once all of the problem blocks were assembled on each form, we prepared the instrument for pilot testing. We wrote a set of instructions with the objective of keeping them short and simple. This set of instructions included an example of exactly how to complete the item. These instructions were nearly identical to those used in the MRO workshops at USASMA, so we had essentially already "tried them out" with NCOs. We then concluded that the measure was ready for pilot testing.

Step 4: Conduct Pilot Test, Key Parallel Test Forms, and Refine Tests

Forms A and B of the LPI were pilot tested with NCOs (n=186) at two different forts. Senior NCOs at USASMA (n=135) also completed the instrument, with their answers providing a key for scoring the LPI. Pilot test and keying sample data were used to refine the forms, producing two parallel forms of 20 items each.

Table 2.9
Number of Problem Blocks by Distribution of Priority Scores

Distribution	Definition	# in Form A	# in Form B
High Value	Problems with priority score above the median	8	8
Constricted	Problems with priority scores within one standard deviation of the mean	8	8
Full Value	Problems across the range of priority scores	8	8
Totals	Entire form	24	24

Once all of the problem blocks were assembled on each form, we prepared the instrument for pilot testing. We wrote a set of instructions with the objective of keeping them short and simple. This set of instructions included an example of exactly how to complete the item. These instructions were nearly identical to those used in the MRO workshops at USASMA, so we had essentially already "tried them out" with NCOs. We then concluded that the measure was ready for pilot testing.

Step 4: Conduct Pilot Test, Key Parallel Test Forms, and Refine Tests

Forms A and B of the LPI were pilot tested with NCOs (n=186) at two different forts. Senior NCOs at USASMA (n=135) also completed the instrument, with their answers providing a key for scoring the LPI. Pilot test and keying sample data were used to refine the forms, producing two parallel forms of 20 items each.

Table 2.10
Problems Contained in Each Item of the Leadership Problems Inventory (Pilot Test Form A)

Item #	A	B	C	D	E
1	A24	B26	B64	B46	A36
2	B53	A44	A24	B29	C45
3	B53	B64	C6	B61	C2
4	A44	A35	A9	B64	C2
5	C30	C45	B61	C49	B26
6	A9	A24	C50	C8	A23
7	B53	C38	B26	A57	C29
8	B23	B26	C50	A25	C30
9	C38	B6	A44	C61	A11
10	B46	A30	A9	C29	C61
11	B51	B46	B6	A33	A22
12	A9	B17	B6	A36	B33
13	A54	C6	C40	A65	C19
14	A9	C25	C61	B28	C27
15	C44	A35	B6	B39	A54
16	C25	A54	A44	B36	A45
17	B28	A65	A54	B10	A48
18	B28	C44	C40	C37	A51
19	A45	C44	A57	A65	A10
20	C40	A45	C50	A48	A65
21	A7	C39	C6	B57	A9
22	B10	C35	B57	B8	A9
23	C37	C64	A8	B4	C6
24	B13	C35	C6	B41	C27

Note: Identification numbers indicate the form letter (A, B, C) and problem number used on the Response Generation Forms.

Table 2.11
Problems Contained in Each Item of the Leadership Problems Inventory (Pilot Test Form B)

Item #	A	B	C	D	E
1	A44	B26	B61	B23	B9
2	C61	C6	A24	A9	A33
3	B26	B53	B6	A35	B29
4	C6	B23	C27	B26	C45
5	A24	A9	B62	C30	B26
6	B53	A24	B61	C29	C12
7	B53	A36	B64	C50	C49
8	B64	A23	B61	A57	C8
9	A22	A44	A23	C45	A9
10	B62	B51	B33	C61	A23
11	B46	C29	A8	B17	A44
12	A11	B51	A44	A25	A33
13	A54	B28	A65	C64	A45
14	C44	C6	B39	B28	C39
15	A65	C26	A44	C40	C44
16	C6	A65	B6	C61	C26
17	A48	C40	A54	A10	C26
18	B28	C61	A57	A48	B57
19	C44	B45	C50	A54	C37
20	C25	A57	B41	A48	A54
21	B45	A9	A35	C6	B13
22	C64	B12	A9	C61	C37
23	C64	C35	A35	A45	A7
24	C61	B4	A45	A35	B8

Note: Identification numbers indicate the form letter (A, B, C) and problem number used on the Response Generation Forms.

[1]One item was unintentionally placed on two of the three forms.

CHAPTER 3

DEVELOP AND EVALUATE CRITERION MEASURES

Overview

The focus of this chapter is the development and initial evaluation of criterion measures of NCO performance. These measures were implemented as part of the validation study. The criterion performance measures tap "typical" NCO job performance as part of a concurrent validation design. Due to the conceptualization of the predictor space, we focused criterion identification on dynamic tasks that require continual monitoring of the task environment with subsequent adjustments in responding. Tasks that we considered for the criterion space were the NCO management and leadership tasks which cannot be preprogrammed but require adaptations of more general rules and schemata to the situation at hand. In this chapter, we briefly review how specific performance dimensions were identified, then describe procedures for developing and revising the criterion measures.

Procedures

We completed a number of activities to develop, evaluate and revise the criterion measures. This process is displayed graphically in Figure 3.1. Initially, our efforts focused on collecting critical incidents to identify the important job dimensions for an NCO in the Army. Then we developed measures to tap NCO job performance on those dimensions. The following activities are described in subsequent sections:

- Develop Performance Dimensions
- Conduct Critical Incident Workshops
- Revise and Finalize Performance Dimensions
- Conduct Retranslation and Instrument Development Workshops
- Draft Criterion Measures
  - Personnel File Form
  - NCO Behaviorally-Anchored Performance Ratings Scales (BARS)
  - Structured Interview

Develop Performance Dimensions

Our first activity was to conduct a literature review to identify performance dimensions relevant to supervisory, managerial and leadership areas. Other literature areas we tapped (organizational citizenship and assessment center research) include relevant Army doctrine and leadership course materials that had information on performance dimensions. From these literature sources, we identified a pool of 281 performance dimensions. We then grouped and sorted these dimensions based on the similarity of the performance content implied in the

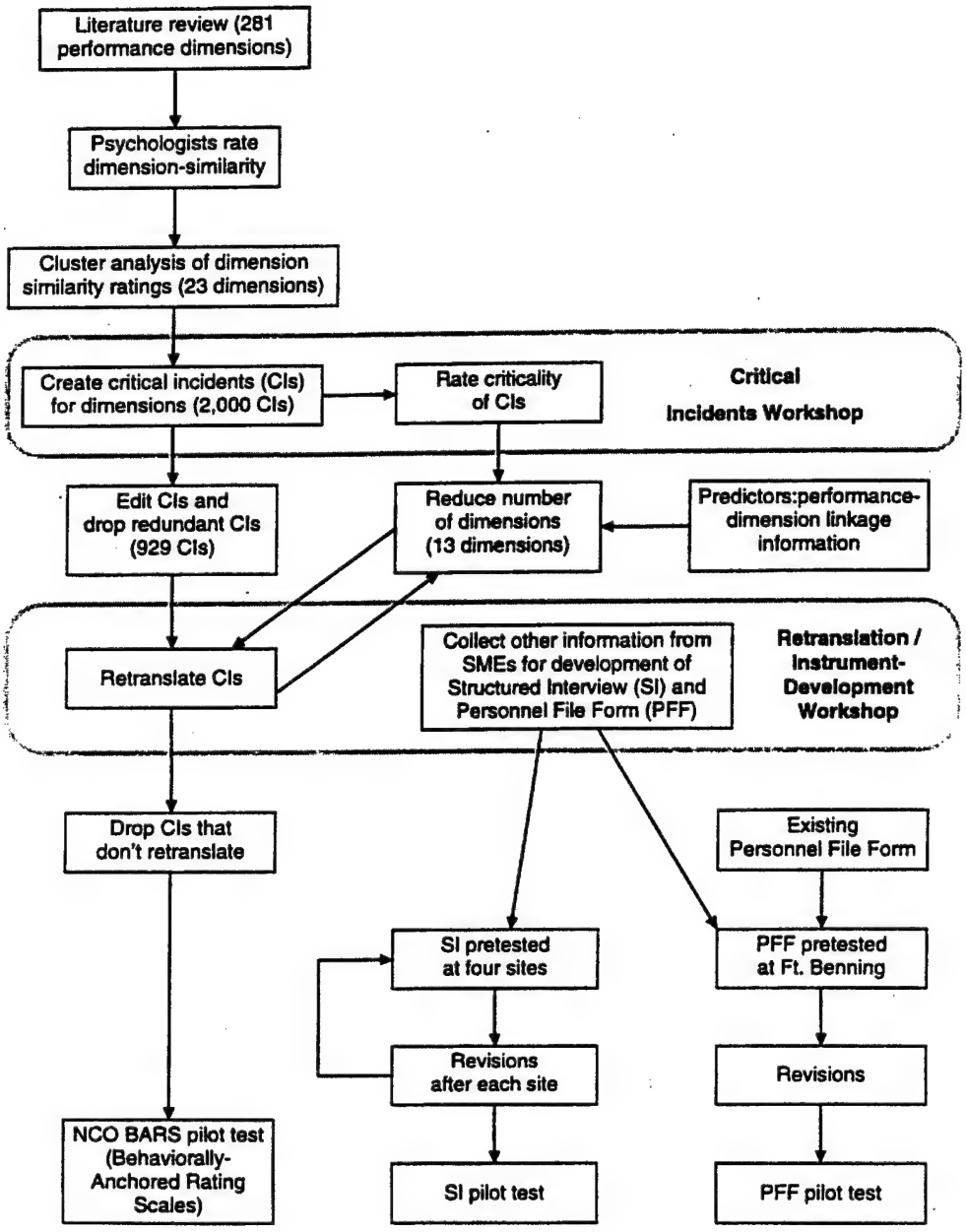


Figure 3.1
Criterion measure development process.

definitions. Eight in-house psychologists formed groupings, and a statistically-based composite solution was achieved by:

- creating a matrix that indicated the number of respondents that put each pair of elements in the same group,
- correlating columns of that matrix, which provided an estimate of the similarity of pairs of elements based on the similarity of how they were grouped with all elements,
- factoring the correlation matrix (placing unities on the diagonal), and
- interpreting clusters of elements based on the similarity of their factor loadings.

Twenty-three clusters were identified and became the proposed criterion dimensions.

Conduct Critical Incident Workshops

We conducted critical incident workshops to determine which of the 23 performance dimensions were most relevant to the Army NCO population. The objectives for conducting these workshops were to: 1) generate performance incidents, 2) collect dimension criticality data, and 3) review existing performance measures (for a review see Peterson et al., 1993).

"Performance" or "critical" incidents are scenarios generated by job experts that describe effective, average, and ineffective performance that they have observed on a job. We planned to use these scenarios to check the accuracy and comprehensiveness of the 23 performance dimensions for the NCO jobs being considered. For example, if Subject Matter Experts (SMEs) were unable to develop incidents for a specific dimension, the dimension may not be relevant for the job. Dimensions for which a large number of incidents are developed are likely to be a major part of the job.

We collected performance incidents following procedures outlined in Pulakos and Borman (1986) and Pulakos (1991). We trained SMEs (in this case, active duty Non-Commissioned Officers) to write incidents, highlighting features of good performance. We also showed SMEs examples of poorly written incidents and instructed them in how to avoid the most common errors in writing incidents. SMEs then wrote a performance incident and were given constructive feedback. Following feedback, SMEs were asked to generate two incidents for each of 23 performance dimensions, one highly effective example and one ineffective example.

SMEs were also asked to rate the amount of time spent and importance of each performance dimension compared to other activities required on their jobs. We calculated the sum of the time spent and the importance ratings for each dimension as a composite index of the criticality of a performance dimension. Then, we used these criticality estimates to identify performance dimensions to be retained, dropped, or merged.

SMEs also discussed any existing performance indicators which would potentially tap each performance dimension. A list was generated to summarize all these suggestions. The

most frequently mentioned suggestions for existing performance measures were well-known administrative indices (e.g., the NCO Efficiency Report [NCOER]).

Over 2,000 critical incidents were collected. We edited critical incidents for spelling, clarity, and format (to specify the situation, the action of the NCO, and the outcome of the incident). We dropped redundant and ambiguous incidents. A total of 929 incidents remained after the editing process. We used these critical incidents to help revise the performance dimensions, develop structured interview questions, and develop the NCO Performance Rating (BARS) scales.

Revise and Finalize Performance Dimensions

Then, through an iterative process, the performance dimensions were revised and finalized. We based initial revisions on critical incident information, criticality ratings of the dimensions, and predictor-performance linkages made by psychologists. Decisions to add, revise, or drop dimensions were based on the answers to the following questions:

1) Were any critical incidents generated for the dimension?
2) Were incidents sorted onto one or more of the dimensions?
3) What were dimension criticality ratings for the dimension?
4) Was the dimension related to new or existing predictors?
5) Did the dimension have similar predictor-performance predicted validities as other dimensions?

The end result of the formalized decision process (see Peterson et al., 1993) was a set of 15 performance dimensions.

Conduct Retranslation and Instrument Development Workshops

The initial activity in preparing for the retranslation exercise was to prepare for the data collection and instrument development workshops. We developed critical incident retranslation packets, developed administrator scripts, submitted a RIRAS, and coordinated with the data collection site and participants.

Develop Critical Incident Retranslation Packets

Approximately 1,000 of the edited critical incidents were assembled into rating packets for distributing to NCOs. The packets were designed to be self-administering, such that NCOs read each incident, categorized it into one of 15 draft performance dimensions, and rated the effectiveness level represented in the incident.

Develop Scripts and RIRAS

Prior to pre-testing the measures and conducting retranslation workshops, project staff developed a set of administrative guidelines to guide the data collection (e.g., scripted

introductions, a project orientation, measure and task descriptions, a background information form). These guidelines, as well as the performance measures (discussed later) and retranslation packets, were submitted to ARI for review in early November, 1993.

Coordinate Data Collection Sites and Participants

The retranslation and pre-test of the measures were coordinated with the predictor development workshops at Fort Benning. Project staff contacted Army personnel at Fort Benning to arrange for 120 workshop participants (NCOs representing levels E-5 through E-8 and a variety of MOS) over a two-day period in early December, 1993.

Critical Incident Retranslation

The NCOs participating in the workshops were first given an introduction to the project and their role in it. They read the Privacy Act statement, filled out a background information form, and then were provided with definitions of each performance rating category and a list of edited critical incidents. They were asked to categorize each critical incident according to the definitions. Participants were to assign each critical incident to the performance category to which it was most similar in content. Participants were also provided with a 1-7 effectiveness scale (1 = not effective, 4 = moderately effective, and 7 = extremely effective) and asked to make effectiveness ratings for each incident. Each NCO participant was asked to retranslate approximately 160 incidents. Six separate retranslation booklets were distributed; each participant completed one booklet.

The retranslation exercise was conducted during four workshops over a two day period (December 14-15, 1993) at Fort Benning. Thirty NCOs were scheduled for each workshop. The workshops were also used to collect information that would facilitate further development and refinement of the three criterion measures for the ECQUIP project: the NCO Performance Rating Scales, the Structured Interview, and the Personnel File Form.

During the workshops each NCO completed one retranslation booklet. A total of 94 NCOs retranslated and rated 929 critical incidents; on average between 15 and 20 NCOs rated each incident.

The 15 performance dimensions were then modified as a result of the critical incident retranslation process. We merged the dimension Analyzing, Problem Solving, and Decision Making with Planning, Organizing, and Coordinating to form Organizing, Coordinating, and Executing; there was a great deal of overlap in the critical incidents which were categorized on the original two dimensions. Additionally, one dimension (Adapting) was dropped altogether because none of the incidents were unambiguously retranslated into the category. A description of the 13 final performance dimensions (listed below) appears in Appendix B.

1) Demonstrating Technical Knowledge and Skill
2) Communicating Orally
3) Writing

4) Demonstrating Effort and Motivation
5) Following Regulations, Policies, and Procedures
6) Demonstrating Integrity and Discipline
7) Relating and Cooperating with Others
8) Motivating Others
9) Planning and Providing for Training
10) Directing, Monitoring, and Supervising Work
11) Organizing, Coordinating, and Executing
12) Demonstrating Responsiveness
13) Representing

Draft Criterion Measures

Initial work on criterion development began during the retranslation workshops at Fort Benning, and continued through the pilot test at Fort Hood. NCO subject matter experts provided input throughout the development cycle. Through an iterative process, we developed three performance measures to assess the performance dimensions. These measures included: 1) a self-report administrative measure (the Personnel File Form), 2) a structured interview, and 3) a set of behaviorally-anchored performance rating scales.

Personnel File Form (PFF)

The initial version of the Personnel File Form was adapted from a form used for collecting administrative information from second-tour soldiers in the Career Force project (Campbell & Zook, 1990). Items on the original PFF that were not useful on the Career Force study were deleted or changed. The PFF was pre-tested at Fort Benning. NCOs also provided PFF data during the workshop where they completed the written predictor materials.

We used data collected at Fort Benning to help revise the Personnel File Form. A total of 69 NCOs completed the draft Personnel File Form. The NCOs also discussed possible modifications and additions to the form. Frequency distributions of the data were examined to identify items with very low or very high levels of endorsement or other problems that would suggest a need for revision. We revised the form to eliminate some items with inherent opportunity bias (e.g., participation in operational deployments) and to clarify items that respondents had difficulty understanding.

Structured Interview

We developed a preliminary set of interview questions from themes which emerged from the critical incidents. For instance, we developed questions to tap Motivating Others because so many critical incidents identified this as an important job dimension. The initial set of questions was pre-tested with a subset of the participants attending the retranslation workshops. This enabled us to determine whether NCOs could respond to the questions adequately. The structured interview was extensively pre-tested at Forts Benning, Eustis, Drum, and Lewis. After the initial pre-test interviews at Fort Benning, we also held discussions with small groups

of interviewed NCOs to obtain their reactions to the questions and any suggestions for revisions and/or other potential questions.

We developed and modified the structured interview through an iterative process of pre-testing (at four sites) and revising. Results from each pre-test site are briefly described next.

Fort Benning. A total of ten NCOs participated in interview development activities at Fort Benning (December 14-15, 1993). The NCOs participated in the structured interview, then provided feedback regarding difficulty answering questions and the relevance of the questions to their jobs.

The structured interview pattern pre-tested at Fort Benning contained many more questions than could be used during an operational interview. Following the pre-test, we combined and eliminated questions on the basis of interviewee responses and comments. Interview responses were also used to provide sample answers on the rating scales. We developed a shorter version of the interview and a set of refined interview rating scales following this pre-test.

Fort Eustis. The data collection at Fort Eustis was completed on January 3-5, 1994. Four NCOs completed the one-on-one structured interview in audio tape-recorded sessions. Four researchers later reviewed and scored the tapes using the draft interview rating scales. We revised the structured interview questions and scales on the basis of the quality of the answers elicited by the interview questions and the consistency of the subsequent ratings.

Fort Drum. The data collection at Fort Drum was completed on January 18-20, 1994, as a part of a larger predictor data collection. Three NCOs completed the structured interview in audio tape-recorded sessions. These interviews were also reviewed by project researchers and scored using the draft interview rating scales. Again, we revised the structured interview questions and scales on the basis of the quality of the answers elicited by the interview questions and the consistency of the subsequent ratings.

Fort Lewis. The data collection at Fort Lewis was completed on April 18-22, 1994. We conducted more extensive pilot testing with the structured interview. Of the total of 23 interviews, 19 were conducted with an additional rater present as a check on interrater reliability. Structured interview questions were revised to elicit desired behaviors and experiences. Interview rating scales were also revised based on the results of this pre-test.

NCO Behaviorally-Anchored Performance Rating Scales (NCO BARS)

We developed the draft version of the NCO BARS by first identifying critical incidents that were unambiguously retranslated. Generally, we retained incidents for use in scale development if greater than 60 percent of the NCOs sorting an incident placed it into a single dimension, and if the effectiveness ratings on the incident varied by less than 1.5 scale points on the 7-point scale. Incidents that met these criteria were used to develop behavioral anchors for each rating dimension.

Experienced project staff wrote behavioral summary statements to capture the content of the specific examples at each of three performance levels (low, average, and high) for each rating dimension. Where too few incidents were available for a category, incidents that initially did not meet the inclusion criteria were modified to clarify language, and were placed on the draft scales for review during the pilot test. We pre-tested the Performance Rating Scales during the pilot test at Fort Hood.

Summary of Criterion Measures and Performance Dimensions

In summary, we developed three criterion measures for use in the validation study. Table 3.1 shows which NCO performance dimensions were assessed by the three measures.

Table 3.1
Performance Dimensions Assessed by Criterion Measures

Performance Dimensions Assessed	Criterion Measure		
	Supervisor Ratings	Structured Interview	Personnel File Form
1. Demonstrating Technical Knowledge and Skill	✓		
2. Communicating Orally	✓	✓	
3. Writing	✓		
4. Demonstrating Effort and Initiative	✓	✓	✓
5. Following Regulation, Policies, and Procedures	✓		
6. Demonstrating Integrity and Discipline	✓		✓
7. Relating and Cooperating with Others	✓		
8. Motivating Others	✓	✓	
9. Planning and Providing for Training	✓	✓	
10. Directing, Monitoring and Supervising Work	✓		
11. Organizing, Coordinating and Executing Work	✓	✓	
12. Demonstrating Responsiveness	✓		
13. Representing the Army	✓		

CHAPTER 4

PILOT TEST DRAFT PREDICTOR AND CRITERION MEASURES

We scheduled three pilot tests at different sites to suit different purposes. Two were scheduled for military installations (Fort Stewart and Fort Hood) to test NCOs similar to those to be included in the validation sample. A third was scheduled with undergraduate students. The purpose of the first pilot test (Fort Stewart) was to try out the new tests to determine timing and format revisions, and to determine estimates of overall time for these tests. We also planned to ask for feedback on the tests in order to identify any typographical errors, unclear instructions or other wording, etc. There was a one-month interval after this pilot test that allowed us to refine our procedures and instruments before traveling to the next military installation. The purpose of the second pilot test (Fort Hood) was to conduct a dress rehearsal, or "tryout", of the procedures for collecting both predictor and criterion data in the future full-scale data collection phase of the validation study.

The third pilot test included university students, enabling us to use a contrasting group's approach to compare responses on the tacit leadership judgment tests to those of NCOs. Having university students as pilot test subjects was a valuable step in evaluating the two new leadership tests.

In this chapter, we describe the NCO and undergraduate pilot tests. For each series of pilot tests, we describe the predictor and criterion measures as they existed at the time of the pilot test, the procedures for data collection, and the pilot test results.

NCO Pilot Tests

Description of Instruments Description of Instruments

Marker Instruments

Verbal Ability: Word Knowledge Test. The Word Knowledge Test from the Army Classification Battery (ACB) (Schratz & Ree, 1989) was used as the Verbal Ability measure for NCOs. There are 20 items on the test and 20 minutes are allowed for completing them. The examinee's test score is the total number correct. For each item on this instrument, the examinee must choose one word from among four alternatives that is most similar to a given word.

Quantitative Ability: Arithmetic Reasoning Test Ability:. Also taken from the ACB, the Arithmetic Reasoning test is a test of an examinee's ability to solve simple arithmetic problems. For each item, the examinee is presented with a brief word problem that requires some numeric calculations. The examinee must then choose from four alternatives. Scratch paper is provided for figuring answers.

Spatial Ability: Assembling Objects. We used a shortened, 20-item version of the Assembling Objects Test that contained ten of each of two types of problems. For each problem, the examinee must figure out how the given object will look when its parts are put together correctly. The reliability of the original Assembling Objects Test is quite high ($r > .90$) (Peterson, Russell, Hallam, Hough, Owens-Kurtz, Gialluca, & Kerwin, 1990; Peterson, Hough, Dunnette, Rosse, Houston, Toquam, & Wing, 1990), allowing us to drop some items to shorten it from the original 36 items down to 20 items, and still maintain a respectable reliability. Ten minutes are allowed to complete this version of the test. The examinee's test score is the total number correct.

Biodata Inventory: Assessment of Background and Life Experiences (ABLE). The Assessment of Background and Life Experiences (ABLE) is a shortened (133 item) version of the instrument developed and used for the Army during the Project A and Career Force projects. It is administered without a time limit. Of the total 133 items, 114 map onto seven substantive scales as follows:

Scale Name	No. Items
1. Work Orientation	28 items
2. Adjustment	15 items
3. Dominance	19 items
4. Dependency	21 items
5. Locus of Control	13 items
6. Cooperation	10 items
7. Physical	8 items

Each item on the seven substantive scales is scored on a three-point basis, and the examinee's score for a given scale is the total of the scores for each item on that scale.

Two additional scales are included on this instrument to detect respondents who: a) do not read and answer items with care (the Non- Random Response scale), and/or b) respond in an overly socially desirable manner (Unlikely Virtues scale). The Unlikely Virtues Scale has 11 items that are also scored on a three-point basis and summed to form an Unlikely Virtues Scale score. In contrast, the 8 items on the Non-Random Response Scale are scored dichotomously, such that the score on the scale is computed as the total number of responses that match the keyed response.

New Measures

Tacit Knowledge: Army Leadership. The ALQ was intended to tap constructs identified in the Base Year Final Report (Peterson et al., 1993) as potentially useful predictors of NCO performance, including NCO Tacit Knowledge, Problem Understanding, and Problem Solving. Each item on the instrument presents a leadership problem scenario along with five alternative actions for handling the problem. Each item requires an examinee to select a "most effective" alternative for handling the problem and also to select a "least effective" alternative for handling

the problem. Because an examinee provides these two responses for every item (instead of the usual "best" response), the instrument potentially allows us to obtain twice as much information from each item. Two different forms (Form A and Form B) of the ALQ, each with 48 items, were prepared for the pilot test.

To assist us in scoring the ALQ, we used the USASMA keying sample statistics. As highly experienced current and former NCOs, students at USASMA form an excellent sample for expert keying. A total of 135 soldiers at USASMA completed the ALQ. The statistic of most interest in developing and using our scoring algorithms was the percentage of USASMA respondents who endorsed each response. In addition to using these statistics to score the ALQ, we also used them to identify items to delete from the final instrument forms to be used in the concurrent validation phase of the project. Our instrument refinement procedures are described in detail in the next chapter.

We developed four alternative algorithms for scoring each item on the ALQ. Each algorithm was applied to the two types of effectiveness responses collected for each item: identification of 1) the Most Effective Response, and 2) the Least Effective response. Each algorithm called for computing a Total score for an item by summing scores for the Least Effective Response and the Most Effective Response. In addition, each of these algorithms allowed for computation of test level scores. Test Most Effective, test Least Effective, and test Total scores were calculated by summing their respective item scores. A description of how each of these algorithms was applied at the item level is shown below.

ALQ Scoring Algorithm 1. Computations for this algorithm are based on the notion of directly comparing the levels of Most Effective and Least Effective endorsement by USASMA examinees for each item response option. Examinees who did not respond to a given item were assigned a score of 0. This algorithm assigned a score on a given item that proportionally matched the attractiveness of the option to the subject matter experts, and hence allowed for consideration of the full manner in which responses to a problem varied in their effectiveness.

Step 1: Identify the option that the examinee endorsed as Most Effective.

Step 2: Compute the difference between the percentage of USASMA examinees who also endorsed the examinee's Most Effective option and the percentage of USASMA examinees who endorsed that same option as Least Effective. This was the Most Effective score for that examinee on that item.

Step 3: Identify the option that the examinee endorsed as Least Effective.

Step 4: Compute the difference between the percentage of USASMA examinees who also endorsed the examinee's Least Effective option and the percentage of USASMA examinees who endorsed that same option as Most Effective. This was the Least Effective score for that examinee on that item.

Step 5: Assign a score of 0 for the Least Effective or Most Effective score if the examinee did not endorse an option for the respective response type.

Step 6: Add the Most Effective score for an item to the Least Effective score for an item to provide the Total score for an item.

ALQ Scoring Algorithm 2. This algorithm assigned a score on a given item that proportionally matched the attractiveness of the option to the experts. The value of 20 was assigned for a "no response" because it equaled the mean percentage computed across all potential responses. This approach assigned a value equivalent to guessing (one in five chance of guessing correctly) rather than penalizing the participant for skipping an item. This added a level of complexity to the algorithm.

Step 1: Compute the percentage of USASMA keying study participants who endorsed each option as Most Effective.

Step 2: Assign the percentage from Step 1 for the option that the examinee endorsed as Most Effective. This was the Most Effective score for that examinee on that item.

Step 3: Compute the percentage of USASMA keying study participants who endorsed each option as Least Effective.

Step 4: Assign the percentage from Step 3 for the option that the examinee endorsed as Least Effective. This was the Least Effective score for that examinee on that item.

Step 5: Assign a score of 20 for the Least Effective or Most Effective score if the examinee did not endorse an option for the respective response type.

Step 6: Add the Most Effective score for an item to the Least Effective score for an item to provide the Total score for an item.

ALQ Scoring Algorithm 3. This algorithm assigned a point for selecting the "correct responses(s)" according to those most highly endorsed in the USASMA keying study. This algorithm is commonly used with situational judgement tests. It is also simple to apply. Unlike other algorithms discussed here, it did not differentiate between the incorrect options in terms of their degree of "correctness."

Step 1: Identify as the "correct Most Effective Response" those option(s) that were most frequently endorsed by USASMA keying participants as Most Effective.

Step 2: Assign one point if the examinee identified the correct Most Effective option. Assign zero points if the examinee marked some other option, or if the Most Effective portion of the item was not answered.

Step 3: Identify as the "correct Least Effective response" those option(s) that were most frequently endorsed by USASMA keying participants as Least Effective.

Step 4: Assign one point if the examinee identified the correct Least Effective option. Assign zero points if the examinee marked some other option, or if the Least Effective portion of the item was not answered.

Step 5: Add the Most Effective score for an item to the Least Effective score for an item to provide the Total score for an item.

ALQ Scoring Algorithm 4. This algorithm assigned a point for selecting the "correct responses(s)" according to those most frequently endorsed in the USASMA keying study, and also penalized a point for selecting the option keyed as the correct response for the opposite scale. Therefore, a penalty point was taken if an examinee (a) selected a Least Effective keyed option as the Most Effective Response or (b) selected a Most Effective keyed option as the Least Effective response. The assumption underlying this scoring algorithm is that it is more incorrect to choose the oppositely keyed response than a response that was not keyed as either Most or Least Effective. Thus, the algorithm applied the traditional method used in #3, with the addition of a penalty. Such a scoring scheme was possible because, unlike other situational judgment test efforts, we had collected both Most Effective and Least Effective responses from the keying sample.

Step 1: Identify as the "correct Most Effective Response" those option(s) that were most frequently endorsed by USASMA keying participants as Most Effective.

Step 2: Identify as the "correct Least Effective response" those option(s) that were most frequently endorsed by USASMA keying participants as Least Effective.

Step 3: Assign one point if the examinee identified the correct Most Effective option. Assign negative one point if the examinee marked the option identified as the correct Least Effective response. Assign zero points if the examinee marked some other option, or if the Most Effective portion of the item was not answered.

Step 4: Assign one point if the examinee identified the correct Least Effective option. Assign negative one point if the examinee marked the option identified as the correct Most Effective Response. Assign zero points if the examinee marked some other option, or if the Least Effective portion of the item was not answered.

Step 5: Add the Most Effective score for an item to the Least Effective score for an item to provide the Total score for an item.

Self-Efficacy: Army Leadership Questionnaire-. Self-Efficacy is also measured via the ALQ. For each problem presented in the ALQ, the examinee indicates (on a 10-point graphic scale) his/her degree of confidence in selecting the "most effective solution." This question is presented after each situational judgment item. Examinees then answer a similar question for

the "least effective solution." These questions are included for every problem on both forms of the ALQ. As there are 48 problems on each form, this means that an examinee completing the ALQ will provide 96 responses relevant to self-efficacy.

Using traditional self-efficacy scoring methods, an instrument score for self-efficacy could be obtained by summing the ratings for each item. As the ALQ asks about confidence ratings on both Least Effective and Most Effective aspects of each item, scale scores for both of these can be calculated. In addition, a total score can be derived by summing the two scale scores. With each of these scores, the higher the mean rating, the greater the self-efficacy score. Throughout pilot testing, we continued to consider other options for scoring self-efficacy.

Priority Setting: Leadership Problems Inventory (LPI)Setting:. The LPI was intended to tap constructs identified in the Base Year Final Report (Peterson, et al., 1993) as potentially useful predictors of NCO performance, including planning, and problem understanding. For each item a respondent reviews a set of five problems (a "problem block") and determines which problem he or she would act on first, second, third, and so on. As such, five responses are obtained for each item. Two forms of the LPI were prepared, each with 24 items.

At the pilot testing stage, various methods for scoring the LPI were considered. The algorithms are based on comparisons between the examinee's responses and those of the subject matter experts in the USASMA keying sample. A total of 135 senior NCOs who were students at USASMA completed either Form A or Form B of the LPI. Each algorithm compares an examinee's responses to the mean ranking computed from the USASMA keying sample for each option. For each algorithm, a total score on the instrument is gained by summing all of the item scores. A description of how each algorithm was applied at the item level is shown below.

Algorithm 1: D^2. This algorithm provides the summed squared difference between USASMA mean rankings and examinee rankings. The ranking applied to an option by the examinee is subtracted from the mean ranking assigned by subject matter experts for that option. This difference is squared and summed with the squared differences for other options on that item. It indicates the degree to which all five responses to an item vary from the responses of the USASMA keying sample. This algorithm was meant to consider the relative priority of all five problems when comparing examinee responses to those of the subject matter experts.

- Step 1: For each option, subtract the subject's rank from the mean rank given by the USASMA SMEs, and square the difference.

- Step 2: For each item, sum the squared differences across options.

Algorithm 2: Low Priority Sum. This method provides the sum of the USASMA mean rankings for the options ranked fourth and fifth priority by the examinee. This algorithm was meant to indicate the amount of agreement between the examinee and the experts on low priority problems. A high score here indicates that the problems ranked fourth and fifth by the examinee were ranked similarly by experts. We used the combined mean for the fourth and fifth ranked

options, because we wanted to focus on general agreement on low priority items, rather than the agreement on only the lowest priority item.

- Step 1: Compute the USASMA mean rank for each response option.
- Step 2: Add the USASMA mean rank for the response option that the examinee ranked 5th to the USASMA mean rank for the response option that the examinee ranked 4th.

Algorithm 3: High Priority Sum. This method provides the sum of the USASMA mean rankings for the options ranked first and second priority by the examinee. This algorithm was meant to indicate the amount of agreement between the examinee and the experts on high priority problems. Note that for this algorithm, a low score (rather than a high score) indicates that the problems ranked first and second by the examinee were ranked similarly by experts. We used the combined mean for the first and second ranked options, because we wanted to focus on general agreement on high priority items, rather than the agreement on only the highest priority item.

- Step 1: Compute the USASMA mean rank for each response option.
- Step 2: Add the USASMA mean rank for the response option that the examinee ranked 1st to the USASMA mean rank for the response option that the examinee ranked 2nd.

Algorithm 4: Difference between Low and High Priority Sums. This method provides the difference between the item scores derived with Algorithms 2 and 3. As stated above, Algorithm 2 focuses on the examinee-expert agreement on low priority problems, while Algorithm 3 focuses on the examinee-expert agreement on high priority problems. This algorithm is essentially the sum of the two methods, so that it estimates the amount of examinee-expert agreement on both high and low priority problems.

- Step 1: Compute the score for Algorithm #2.
- Step 2: Compute the score for Algorithm #3.
- Step 3: Subtract the score for Algorithm #3 from that for Algorithm #2.

Examinees are to rank all options for each item, and are therefore to enter each option only once on the ranking list for an item. If an examinee lists an option more than once for a given item, all five responses were treated as missing (regardless of the algorithm used).

Criterion Measures

In the following sections, we provide a description of the criterion measures that we administered to NCOs during the pilot test at two sites, Forts Stewart and Hood.

Personnel File Form. The Personnel File Form is a self-report instrument which asks respondents a total of 11 questions, tapping five performance areas. The questions tap Awards/Badges/Commendations, Education, Promotions, Test Results, and Disciplinary Actions.

Scores for the Personnel File Form can be calculated a number of ways. For example, a method for weighting or averaging the different sections could provide an overall score for each individual. Alternatively, scoring rules for "scale" scores have also been considered for final data analyses. The procedure for developing a final score (or scores) on the PFF has not yet been determined.

NCO Behaviorally Anchored Rating Scales. The supervisor rating form is a measure of supervisor ratings of subordinates along 13 dimensions of NCO performance. For each dimension, a definition is provided along with a description of low, moderate and high performance. Also included in the rating scales are specific behavioral examples of low, moderate and high performance (drawn from critical incident information).

There are also eight scales for making Situation Performance Ratings that measure how well the NCO handles certain types of situations on the job. These rating scales tap concrete job behaviors which the supervisor will likely have observed the NCO performing. The supervisors are provided a packet of the rating scales and verbal instructions for completing the ratings. The performance dimensions are described, along with typical rating errors. These performance dimensions are based on the eight categories derived from the Problem Solving Workshop discussed in Chapter 2. The NCO performance rating scales provide a score (from 1-7) for each of the performance dimensions for each NCO.

Structured Interview. The structured interview taps behaviors and situations the individual has experienced. The interview includes standardized instructions which are read to the participant prior to beginning the questions. The interviewer is instructed to probe for complete information for each question.

The interview consists of nine questions which tap five performance dimensions; Motivating Others, Demonstrating Effort, Planning and Providing for Training, Organizing, Coordinating and Executing, and Communicating Orally. The interview takes approximately 40 minutes to complete. When the interview is finished, the interviewer makes ratings along the five dimensions. Scores on individual questions are averaged to derive one score for each of the five performance dimensions listed above.

We videotaped the interviews that were conducted during the pilot test. This allowed us to assess the reliability of interviewer ratings. It also provided us with material to train interviewers for the validation study.

Job Analysis Questionnaire. The Job Analysis Questionnaire describes the 13 job performance dimensions. The respondent is asked to make two ratings (each on a six-point scale) for each of the dimensions. The respondent rates the extent to which the activity takes up their time and attention. They also rate the extent to which successfully performing each activity is important to the success of their unit's mission. The instrument was used to verify the importance of the dimensions for NCOs in a variety of positions and to investigate the similarity of NCO positions.

Procedures

We completed a number of activities to prepare for collecting pilot test data. These activities included meeting with the project sponsors to review progress to date and make decisions about the data collection, preparing materials, and coordinating with the data collection sites. These activities are described below.

Review and Plan With Project Sponsors

In preparation for the pilot tests, we conducted a meeting with ARI project monitors on June 23, 1994 to review the tests we were preparing. We discussed the project progress, reviewed the data collection schedule for the next scheduled trips, and provided examples of the two types of tests we had developed (i.e., the Army Leadership Questionnaire and the Leadership Problems Inventory). In addition, we discussed measures to use as markers of verbal and quantitative ability, specifically, rather than using ASVAB tests, ARI monitors suggested using the Arithmetic Reasoning and Word Knowledge tests from the Army Classification Battery (ACB). Use of the ACB would allow us to measure the same constructs with the same item type, and avoid problems associated with using an operational test.

We established the test order, set item totals and times for timed tests, and decided on preliminary guidelines for time allowances on untimed measures, as shown in Table 4.1.

Table 4.1

Test Order and Times for Pilot Test

Test Name	Type of Measure	Testing Time
Introduction/Background Information Form	Untimed	15 min
Word Knowledge (Verbal Ability), 20 items	Timed	6 min
Arithmetic Reasoning (Quantitative Ability), 20 items	Timed	20 min
Assembling Objects (Spatial Ability), 20 items	Timed	10 min
Army Leadership Questionnaire, 48 items	Untimed	60 min
Leadership Problems Inventory, 24 items	Untimed	60 min
Job Analysis Questionnaire, 13 items, 2 ratings each	Untimed	5 min
Personnel File Form*, 11 items	Untimed	10 min
Assessment of Background and Life Experiences, 133 items	Untimed	25 min
		181 min plus admin time of 10 min = 191 min total

*Note. The Personnel File Form was prepared for administration only at Fort Hood, not at Fort Stewart.

Develop Predictor Materials and Script for Pilot Tests

We prepared two alternate formats for administering the new supervisory judgment tests at Fort Stewart -- one format used separate answer sheets and the second included space in the booklets for writing answers. We planned to use the results of this preliminary test for deciding which response format to use for the larger pilot test at Fort Hood.

We prepared the following materials for administration:

- Background Information Form
- Word Knowledge Test
- Arithmetic Reasoning Test
- Assembling Objects Test
- Army Leadership Questionnaire
- Leadership Problems Inventory
- Job Activities Questionnaire
- Assessment of Background and Life Experiences

Additionally, the following materials were prepared for the pilot tests:

- a script for introducing the project purpose and background
- an agenda
- the Privacy Act
- answer sheets for all tests administered in booklet format

Develop Criterion Materials and Procedures

We reviewed and made minor revisions to instruments that had been developed and revised over the course of the earlier data collections. We also assembled materials necessary for the data collection (e.g., forms, pencils, video camera, videotape, etc.).

Structured Interview. We developed procedures for scheduling interview participants and data collectors for each exercise such that interviewer pairs were rotated frequently so each interviewer appeared several times on videotape. The structured interviews were scheduled to be conducted with two interviewers, allowing for an examination of interrater agreement. Each interview was scheduled for 40 minutes, leaving additional time for ratings to be compiled and discussed immediately following each interview. Additional time was also allowed for discussion of interview questions and the interview scales.

In a half-day group session, all interviewers were trained to conduct interviews. Trainees read interview materials prior to training. Training included how to open the interview, gather information in the appropriate response format (Situation, Action, Outcome), ask probes and follow-up questions, and evaluate the interviewee. Trainees practiced with probing questions and rated actual interviews (presented on audiotape) as part of the training.

Personnel File Form (PFF) and Job Analysis Questionnaire (JAQ). The PFF and JAQ were administered along with the predictor information.

NCO BARS. The pilot test was scheduled to be conducted in two phases: first, participants were asked to complete the rating scales by rating a subordinate NCO; second, a group discussion regarding the scale content was led by a project researcher. The discussion focused on the clarity of the instructions, the appropriateness of the dimensions and the definitions for rating NCOs, the appropriateness of the anchors for each scale, and the usefulness of the behavioral examples for each scale. Particular attention was focused on behavioral examples that were derived from critical incidents that had been modified by project researchers prior to their placement on the scale.

Coordinate Data Collection Sites and Participants

As mentioned earlier, pilot tests were scheduled at Fort Stewart and Fort Hood. At Fort Stewart, arrangements were made to have 90 NCOs representing a variety of MOS and the three rank groups (E-5, E-6, and E-7/8) attend one of six workshops scheduled over a three-day period. At Fort Hood, we requested 240 NCOs (meeting the same specifications) over a four-day period. NCOs from pay grades E4 through E8 participated in the structured interview, while only NCOs in pay grades E6 through E8 were selected for pilot testing of the NCO BARS.

Preparation for pilot testing involved organizing information that had been collected at each of the pre-test sites, completing any additional revisions to the instruments, training data collectors and interviewers, and developing data collection procedures.

Conduct Pilot Test Sessions

The pilot tests were conducted at Fort Stewart on July 13 - 15, 1994, and at Fort Hood on August 1 - 5, 1994. At each session, we described the project background and the NCOs' role in the pilot testing phase of the project, passed out the Privacy Act, and asked participants to start by filling in the Background Information Form. We conducted the timed marker tests (verbal, quantitative, and spatial) first, then took a break before conducting the untimed measures. We handed out the remaining measures, reviewed the instructions for the Army Leadership Questionnaire, and told subjects to continue (and to ask questions as necessary) to complete the materials.

In order to verify that our time limits for each of the timed tests were reasonable, we closely monitored when the subjects actually finished each of the tests. Further, to get an idea of reasonable time limits to allow for the untimed measures, we recorded how long NCOs took to complete each of these instruments.

At the end of each session, we asked NCOs for both general feedback and specific comments about the new leadership judgment tests. We recorded the general comments and

suggestions they made and invited them to write comments pertaining to specific problems in the appropriate spots in their booklets.

The NCO sample consisted of 22 females and 226 males, with 8 missing data for gender. The mean tenure in the Army was 10 years and 5 months. The breakout for race is reported in Table 4.2.

Table 4.2
NCO Pilot Test Sample: Race Frequency and Percent of Sample

Race	Frequency	Percent
White	111	43.4
Native Am./Am. Indian	2	.8
Hispanic	28	10.9
African-American	101	39.5
Asian	1	.4
Other	3	1.2
Missing	10	3.9
Totals	256	100

Note: Race was missing for remainder of sample.

The ranks for the 70 Fort Stewart NCOs and the 186 Fort Hood NCOs who participated in the pilot test are reported in Table 4.3.

Marker Tests

We conducted a similar set of analyses for each of the four marker tests. First, we conducted an item analysis for each test. Next, we looked at the score distributions, computed descriptive statistics, and conducted a supplemental analysis to determine the numbers of missing items across all respondents. Descriptive statistics for these four tests are given in Table 4.4. Finally, we computed the intercorrelations between each of these tests in the NCO pilot test. These correlations are shown in Table 4.5.

Table 4.3
NCO Pilot Test Sample: Rank Frequency and Percent of Sample

Rank	Frequency	Percent
E-3 (promotable)	1	.4
E-4	18	7.0
E-5	124	48.4
E-6	58	22.7
E-7	33	12.9
E-8	4	1.6
Missing	18	7.0
Totals	256	100.0

Table 4.4

Descriptive Statistics for Marker Tests, NCO Pilot Test Sample

Marker Test	n	Mean	S.D.	Hoyt r[1]
Word Knowledge	238	13.37	3.21	.73
Arithmetic Reasoning	237	11.59	4.08	.81
Assembling Objects	234	13.67	4.14	.83
ABLE				
Work Orientation	172	69.81	7.75	.87
Adjustment	172	34.84	4.95	.80
Dominance	172	47.50	5.11	.80
Dependability	172	50.71	5.70	.77
Locus of Control	172	32.76	4.51	.82
Cooperation	172	22.37	2.45	.51
Physical Condition	172	18.58	3.20	.77
Unlikely Virtues	172	16.39	3.13	.63
Non-Random Response	172	7.27	1.41	.77

Table 4.5

Correlation Matrix of Marker Tests, NCO Pilot Test Sample

Marker Test	1	2	3	4	5	6	7	8	9	10	11	12
1. Word Knowledge		..										
2. Arithmetic Reasoning	50	..										
3. Assembling Objects ABLE	27	49	..									
4. Work Orientation	12	17	06	..								
5. Adjustment	12	17	12	37	..							
6. Dominance	19	15	06	60	34	..						
7. Dependability		-05	-03	-07	32	26	16	..				
8. Locus of Control	12	20	14	50	42	33	31	..				
9. Cooperation	08	09	-07	39	02	34	03	10	..			
10. Physical Condition	-11	-03	-08	21	13	28	01	02	11	..		
11. Unlikely Virtues	31	-04	-01	22	14	05	24	18	00	05	..	
12. Non-Random Response	09	08	23	31	20	20	24	16	12	-05	-14	..

Note. Decimal points omitted; ".." = 1.0; n = 172.

Results

New Measures

Tacit Knowledge: Army Leadership Questionnaire. NCO pilot test data were analyzed at the instrument level and the item level. Both sets of statistics were used in refining and evaluating the usefulness of the ALQ.

Descriptive statistics and Hoyt and Split-Half reliabilities were calculated for each form and each scoring technique and are presented in Table 4.6. The majority of values for Hoyt reliabilities ranged from .50 to .86 for the four algorithms, for both Forms A and B. Two of the Hoyt reliabilities for scoring algorithm #1 of Form B reported in Table 2.5 are negative, indicating poor reliability. (It is mathematically possible for a sample reliability estimate to be below zero, even though the true reliability cannot be less than 0.) Tables 4.7 and 4.8 contain correlation matrices detailing the relationship among the four scoring techniques for the ALQ Forms A and B, respectively. These descriptive statistics and correlations reveal that the reliability of the ALQ does not vary greatly across scoring techniques, and the relationship among the scores from different algorithms is quite high. Therefore, the remainder of the instrument statistics are provided using the "Total" score associated with Algorithm #4, as this algorithm is relatively simple (see page 4-5 for a description of this scoring algorithm), and has adequate reliability estimates associated with it.

Table 4.6

Descriptive Statistics and Reliabilities Associated with Various Tacit Knowledge ALQ Scoring Techniques for Each Form: NCO Pilot Test Sample

Scoring Algorithm	N	Mean	S.D.	Hoyt r	Split-Half r
		Form A			
Most Effective	126	1153.1	395.2	.75	.78
Least Effective	126	994.1	499.0	.84	.88
Total	126	2147.2	825.1	.86	.89
Algorithm 2					
Most Effective	126	1675.3	264.8	.73	.77
Least Effective	126	1576.1	311.0	.82	.84
Total	126	3251.4	526.8	.86	.89
Algorithm 3					
Most Effective	126	25.7	5.1	.64	.66
Least Effective	126	23.6	6.1	.75	.73
Total	126	49.3	10.2	.81	.82
Algorithm 4					
Most Effective	126	21.8	7.5	.71	.78
Least Effective	126	19.0	9.8	.82	.83
Total	126	40.9	15.8	.84	.86
		Form B			
Algorithm 1					
Most Effective	124	1091.3	351.9	.69	.71
Least Effective	124	1026.0	468.3	a	.83
Total	124	2117.2	757.0	a	.87
Algorithm 2					
Most Effective	124	1619.6	234.3	.66	.68
Least Effective	124	1613.1	315.1	.81	.81
Total	124	3232.7	499.9	.84	.86
Algorithm 3					
Most Effective	124	21.4	4.3	.50	.51
Least Effective	124	21.6	5.8	.73	.72
Total	124	43.0	8.8	.75	.71
Algorithm 4					
Most Effective	124	18.3	6.1	.60	.62
Least Effective	124	18.1	8.1	.77	.76
Total	124	36.3	12.8	.79	.77

Note: a - Reliabilities were negative. It is mathematically possible for a sample reliability estimate to be below zero, even though the true reliability cannot be less than zero.

Table 4.7
Correlations Among Tacit Knowledge ALQ Form A Scores, NCO Pilot Test Sample

Label	1	2	3	4	5	6	7	8	9	10	11	12
Algorithm 1												
1.Most	..											
2.Least	70	..										
3.Total	90	94	..									
Algorithm 2												
4.Most	98	69	89	..								
5.Least	69	98	92	67	..							
6.Total	90	93	99	90	93	..						
Algorithm 3												
7.Most	96	69	87	95	68	88	..					
8.Least	66	96	90	65	98	90	65	..				
9.Total	87	92	98	86	92	98	89	93	..			
Algorithm 4												
10. Most	97	67	87	94	67	87	97	65	88	..		
11. Least	67	98	92	67	97	91	67	98	92	66	..	
12. Total	88	93	98	86	92	98	88	91	99	89	93	..

Note: Decimal points omitted; ".." = 1.0. n = 126.

Table 4.8
Correlations Among Tacit Knowledge ALQ Form B Scores. NCO Pilot Test Sample

Label	1	2	3	4	5	6	7	8	9	10	11	12
Algorithm 1												
1.Most	..											
2.Least	70	..										
3.Total	90	94	..									
Algorithm 2												
4.Most	98	67	87	..								
5.Least	68	99	93	65	..							
6.Total	89	94	99	88	93	..						
Algorithm 3												
7.Most	91	56	77	94	53	77	..					
8.Least	65	96	89	62	96	90	49	..				
9.Total	87	91	97	87	90	97	82	90	..			
Algorithm 4												
10. Most	97	64	85	95	62	83	95	59	85	..		
11. Least	67	98	92	65	97	92	54	98	91	63	..	
12. Total	89	92	98	86	90	97	79	90	98	87	93	..

Note: Decimal points omitted; ".." = 1.0. n = 124

NCO pilot test statistics for each ALQ item are presented in Appendix C. The information presented includes the following statistics for each score on each item:

- N: number of respondents
- Mean: mean score
- S.D.: standard deviation of the scores
- Part-whole correlation: correlation between the score for that item using that algorithm, and that score type (i.e., Most, Least, Total) with the scale score for a given algorithm, and score type. These are presented for every algorithm and score type.

These statistics were important to ensuring that responses varied sufficiently to develop a useful score on the instrument. We also used these statistics in "balancing" the item difficulty across the two final test forms (to be used in the concurrent validation phase).

Self-Efficacy: Army Leadership Questionnaire. Data from two different NCO samples the NCO pilot test and USASMA keying samples were analyzed to inform revision decisions for the Self-Efficacy Scales of the ALQ. Table 4.9 contains descriptive statistics and reliabilities for these scales in both samples. It is interesting to note from these table entries that the reliabilities for the self-efficacy scales are high (.99), indicating either high internal consistency or evidence of halo effects on the self-report measures. However, the self-efficacy mean scores were higher for the more experienced individuals in the keying sample than for the NCOs in our pilot sample. This is consistent with expectations for an effective measure of self-efficacy: individuals with more experience should realistically rate themselves higher in efficacy for handling the types of problems supervisors face. We chose to report effect size here, rather than to conduct and report results of a significance test of the difference between two means. Effect sizes convey more information about relationships than do statistical significance findings, and statistical significance can sometimes be misleading. (Schmidt, 1996).

The next two tables contain correlations showing relationships among Self-Efficacy Scale scores collected in the NCO pilot test. Tables 4.10 and 4.11 provide these statistics for Forms A and B, respectively.

As the correlation between various self-efficacy scales is quite high, it would be redundant to report additional instrument level statistics for all of the different scales. Therefore, the remainder of instrument level statistics that include self-efficacy will use the "total" self-efficacy score.

Leadership Problems Inventory (LPI). NCO pilot test data were analyzed at the instrument level and the item level. Both sets of statistics were used in refining and evaluating the usefulness of the LPI.

Descriptive statistics and Hoyt and Split-Half reliabilities were calculated for each form and each scoring technique, and are represented in Table 4.12. Table 4.13 contains correlation

matrices detailing the relationship among the four scoring techniques for the LPI Forms A and B respectively. These statistics reveal that the reliability of the LPI varies somewhat across scoring techniques, and the relationship among the scores from different algorithms is quite high, except for the D^2 technique. Therefore, the remainder of the instrument statistics are provided using Algorithm #4, as this algorithm has adequate reliability, and the measure appears comprehensive in that it captures item performance at both high and low levels of problem priority.

Table 4.9
Descriptive Statistics for Self-Efficacy Scales Used with the Army Leadership Questionnaire, NCO Pilot Test and USASMA Keying Samples.

	NCO Pilot				USASMA Keying Sample		
Scoring Algorithm	n	Mean	SD	Hoyt r	n	Mean	SD
				Form A			
Most	90	384.3	69.7	.99	72	406.7	68.1
Least	90	363.7	85.4	.99	73	392.5	71.8
Total	90	747.9	141.7	.99	73	799.6	136.9
				Form B			
Most	81	381.8	77.3	.99	62	407.3	60.7
Least	81	354.2	90.7	.99	62	400.2	60.0
Total	81	736.0	156.7	.99	62	807.4	118.7

Table 4.10
Correlations Among Form A Self-Efficacy Scores, NCO Pilot Test Sample

Label	1	2	3
1.Most	..		
2.Least	68	..	
3.Total	90	93	..

Note. Decimal points omitted; ".." = 1.0.

Table 4.11
Correlations Among Form B Self-Efficacy Scores, NCO Pilot Test Sample

Label	1	2	3
1.Most	..		
2.Least	74	..	
3.Total	92	94	..

Note. Decimal points omitted; ".." = 1.0.

Table 4.12
Descriptive Statistics and Reliabilities Associated with Various LPI Scoring Techniques for Each Form: NCO Pilot Test Sample

Scoring Algorithm	n	Mean	S.D.	Hoyt r	Split-Half r
			Form A		
ΣD^2	115	107.7	17.8	0.14	0.35
Low Priority Sum	115	160.9	7.4	0.74	0.76
High Priority Sum	115	125.9	7.9	0.53	0.84
High/Low Diff	115	35.0	15.0	0.79	0.84
			Form B		
ΣD^2	121	105.7	19.7	0.40	0.41
Low Priority Sum	121	158.1	8.9	0.81	0.85
High Priority Sum	121	129.7	8.7	0.61	0.83
High/Low Diff	121	28.4	17.2	0.83	0.88

Table 4.13
Correlations Among LPI Scores, NCO Pilot Test Sample

Label		1	2	3	4	5	6	7	8
LPI-A									
1.	ΣD^2	..							
2.	Low Priority Sum	-50	..						
3.	High Priority Sum	61	-91	..					
4.	High/Low Diff	-57	98	-98	..				
LPI-B									
5.	ΣD^2	--	--	--	--	..			
6.	Low Priority Sum	--	--	--	--	-52	..		
7.	High Priority Sum	--	--	--	--	58	-93	..	
8.	High/Low Diff	--	--	--	--	-56	98	-98	..

Note: Decimal points omitted; ".." = 1.0. Examinees administered Form A of the LPI were not administered Form B and vice-versa.

Pilot test statistics for each LPI item are presented in Appendix D. These statistics summarize the rankings provided by subjects in the pilot tests and compare these rankings with those provided in the USASMA keying sample. The following statistics on each response option within each item are included:

In the mean of the rankings assigned by participants in each sample

- the rank determined from the USASMA keying sample
- the rank when the mean rankings for the sample are placed in ascending order.
- the difference between the rank determined in the USASMA keying sample and the rank determined in the pilot of concern--the sum of these differences is also reported for each item

We investigated these statistics carefully to ensure sufficient response variability to develop a useful score on the instrument. We also used these statistics to assist in balancing item difficulty on the final test forms (to be administered in the concurrent validation phase).

We calculated keying sample statistics for each LPI item. These statistics allowed us to summarize the rankings provided by participants in the USASMA keying sample and compare these rankings with those derived by applying priority scores from the USASMA priority setting workshop. The information reviewed includes the following statistics on each response option within each item:

- the mean of the rankings assigned by USASMA participants,
- the rank determined from the USASMA keying study,
- the rank determined from the USASMA Multiple Rank Ordering Workshop, and
- the difference between the rank determined in the USASMA Multiple Rank Ordering Workshop and the rank determined in the USASMA keying study the sum of these differences is also reported for each item

In addition to using these statistics to develop a scoring key for the LPI, we referred to these statistics when making decisions about which items to retain (those for which there was minimal disagreement between expert groups about the correct answers).

Table 4.14 reports correlations among each of the new predictor instruments and the marker tests administered in the NCO pilot. These correlations reveal that the ALQ and LPI scores have a moderate, positive relationship with scores on the ABLE Non-Random Response scale. In addition, ALQ and LPI scores generally have a low positive relationship with scores on various measures of cognitive ability. This information is useful in determining that scores on the ALQ and LPI are at least systematic persons who attend to the other instruments are also attending to how they complete the ALQ and the LPI. In addition, the correlations with cognitive ability scores are sufficiently low so that there is room for the ALQ to add to the prediction of performance after considering cognitive ability scores. There are moderate to high correlations between the ALQ and LPI, and generally low correlations between the ABLE and all the new measures.

Table 4.14

Correlations Among New Predictor Instruments and Marker Tests, NCO Pilot Test Sample

Label	1	2	3	4	5	6
1. ALQ, Form A	..					
2. ALQ, Form B	--	..				
3. Self-Efficacy, Form A	18	--	..			
4. Self-Efficacy, Form B	--	42	--	..		
5. LPI, Form A	39	63	-07	-01	..	
6. LPI, Form B	42	53	-11	34	--	..
Marker Tests						
Word Knowledge	29	37	51	35	22	20
Arithmetic Reasoning	28	22	27	18	27	24
Assembling Objects	20	26	20	18	01	27
ABLE						
Work Orientation	03	25	05	12	10	27
Adjustment	03	16	21	02	08	01
Dominance	-02	11	20	01	02	06
Dependability	01	17	-23	-02	20	01
Locus of Control	17	20	03	12	12	11
Cooperation	-06	-04	-01	-14	-10	-07
Physical Condition	-08	-10	21	12	-06	-15
Unlikely Virtues	-13	-17	04	-02	08	-26
Non-Random Response	41	33	11	05	35	55

Note. Decimal points omitted; ".." = 1.0. Examinees administered Form A of new instruments were not administered Form B and vice-versa. However, the new instrument forms were administered independently; for example, examinees who had Form A of the ALQ could have taken Form B of the LPI.

Table 4.14 shows that self-efficacy scores have a moderate, positive relationship with scores on Word Knowledge and somewhat lower correlations with the other cognitive tests. Self-efficacy scores have low correlations with the ABLE measures. However, the correlations of Forms A and B of Self-Efficacy and ABLE non-random response are lower (.11 and .05, respectively). Our explanation for this difference is that respondents tended to fall into a response pattern for the self-efficacy items (e.g., endorsing a similar scale point for all items due to the repetitive nature) and stopped attending to them.

Criterion Measures

Personnel File Form. The total sample size that responded to the PFF at Fort Hood was 186 NCOs. Frequency distributions of NCO responses were analyzed to determine the usefulness of the questions.

NCO Behaviorally Anchored Rating Scales. Two groups of six participants reviewed the NCO BARS on the second and third days of the pilot test at Fort Hood. Their suggestions were used to revise the wording of the form.

Structured Interview. Interviews were conducted over three days. Half of the interviews were videotaped in order to develop interviewer training tapes for later use. Generalizability analyses were conducted to determine reliability of the interview questions. Single-rater reliability was .77 for making relative decisions (decisions between NCOs across the 5 dimensions).

Job Analysis Questionnaire. Descriptive statistics were run on the JAQ. They are reported in Table 4.15.

Table 4.15

Descriptive Statistics for the Job Analysis Questionnaire, NCO Pilot Test Sample

	Importance		Time Spent	
Activity	Mean	SD	Mean	SD
Demonstrating Technical Knowledge and Skill	4.11	.81	3.51	1.00
Communicating Orally	4.10	.91	3.52	1.01
Writing	3.51	1.06	2.81	1.23
Demonstrating Effort and Motivation	4.25	.73	3.92	.88
Following Regulations, Policies, and Procedures	4.19	.88	3.87	.97
Demonstrating Integrity and Discipline	4.40	.75	4.03	.89
Relating and Cooperating with Others	4.13	.76	3.80	.89
Motivating Others	4.19	.88	3.75	1.00
Planning and Providing for Training	4.11	.96	3.55	1.10
Directing, Monitoring, and Supervising Work	4.09	1.00	3.72	1.15
Organizing, Coordinating and Executing	3.99	1.02	3.49	1.20
Demonstrating Responsiveness	4.07	.85	3.78	.92
Representing	3.53	1.33	3.13	1.33

Undergraduate Student Pilot Test

The purpose of including a sample of university students in our pilot testing was to gather evidence to determine if the new instruments under development were measuring constructs that were at least somewhat independent of cognitive ability and other factors that traditionally relate to effective test performance. For the new instruments to provide predictive power beyond these traditional measures, the mean performance of a reasonably intelligent sample of individuals, (i.e., undergraduate students from George Mason University who have no training or experience in the Army), should be lower than that of a sample of NCOs. In addition, a significant portion of the variance that is found in student performance should be attributable to traditional factors

that tend to relate to effective test performance. Therefore, we would expect a relatively high relationship in the student population between our new instruments and measures of cognitive ability.

To address these issues, we gathered a set of measures from the students in addition to the new instruments that we are developing for this project. Some of these measures were taken in the form of paper-and-pencil tests, while others came in the form of self-report demographic data.

Description of Instruments

Marker Measures

General Cognitive Ability. We measured this construct and its subconstructs through a paper-and-pencil test and various self-report measures.

We used the Employee Aptitude Survey-Verbal Reasoning (EAS-VR) subtest to examine cognitive ability using a paper-and-pencil test. The EAS-VR is a five-minute test containing logic problems that require an examinee to read a list of four or five statements of fact. Then the examinee must read each of five conclusions and decide whether each is true, false, or uncertain, based on the given facts. There are six sets of facts and conclusions, for a total of 30 potential responses. The score for this test is the total number correct. The test is a short, reliable measure (r=.82) that is well supported by research as an effective measure of reasoning ability (Ruch, Stang, McKillip & Dye, 1994).

We developed a background information form containing 12 basic demographic and biodata items. Some of the items that we expected to reflect general cognitive ability included self-report queries regarding:

- Grade Point Average (GPA)
- Scholastic Aptitude Test (SAT) scores
  - verbal
  - quantitative

Students were asked to report their undergraduate GPA, and high school GPA if they did not yet have an undergraduate GPA established.

Experience. The background information form also included items meant to tap the experience of the students. It was thought that various types of experience should be related to performance on the new instruments. The experience items asked the students to self report their:

- experience in supervisory positions within a

- job
- club or social organization
- knowledge of/familiarity with the military system and language
- military experience

Supervisory experience items asked whether, and for how long, the respondent had experience in supervisory positions. One item asked for students to rate, using a seven-point scale, their degree of familiarity with the military life, system, and jargon. The follow-up question for those who responded that they had some degree of familiarity (e.g., a rating of "2" or greater) asked how that familiarity was gained (e.g., military family or friends).

Social Skills. The "Social Skills Background Data Measure" is a 30-item instrument developed by Zaccaro and his colleagues at George Mason University for use with undergraduate students with little or no work experience. The instrument assesses a respondent's methods of dealing with a variety of social situations, and has been used to investigate the relationship between social skills and leader performance (Zaccaro, Gilbert, Thor, & Mumford, 1991; Zaccaro, Gilbert, Marks, Connelly, and Mumford, in preparation). The measure has five subscales, including Tolerance for Ambiguity, Dominance, Social Adroitness, Social Perceptiveness, and Oral Expression. A maximum of 35 minutes is allowed for administration. Each question has five response options; each set of response options is graduated in terms of amount or frequency, either from high to low or vice versa. The score on each subscale is the mean of the scores on each item. We used this social skills measure with the student sample so that we could look at the relationships among social skills and scores on the new measures of prioritization and handling leadership situations.

Demographic Information. To help us to better understand our student sample, and to provide information that might be used to examine various hypotheses regarding moderator variable hypotheses, we also asked the students to report their:

- date of birth,
- gender,
- race, and
- undergraduate major.

New Measures

Tacit Knowledge: Army Leadership Questionnaire. We used the same version of the ALQ as prepared for the NCO pilot test.

Self-Efficacy: Army Leadership Questionnaire. We used the same version of the self-efficacy measures for students that we prepared for the NCO pilot test.

Priority Setting: Leadership Problems Inventory. We used the same version of the LPI for students that we prepared for the NCO pilot test.

Procedures

For this pilot test, we prepared materials for recruiting student subjects and piloting the instruments. Arrangements were also made with graduate students to recruit the undergraduate students and run the sessions.

We prepared the following materials for the undergraduate student pilot:

- Background Information Form
- Employee Aptitude Survey
- Army Leadership Questionnaire
- Leadership Problems Inventory
- Social Skills Background

Also, similar to the NCO pilot tests, the following materials were prepared for the undergraduate student pilot:

- a script for introducing the project purpose and background
- an agenda
- the Privacy Act
- answer sheets for all tests administered in booklet format

Undergraduate psychology students were recruited from two summer sessions and one fall session. Students were tested in several sessions between June 25 and September 30, 1994. In the sessions, which ranged in size from five to twenty students, consent to participate was obtained, then the project and the students' role in the pilot test was explained. Students completed the timed test (the Employee Aptitude Survey) first, followed by the untimed measures, (in the following order: Background Information Form, ALQ, Social Skills, LPI).

A total of 120 students were tested. The majority of students ranged in age from 18 - 21 years, with a small percentage of older students. Of the 120 total, 38 (31.7%) were male and 82 (68.3%) were female.

Results

We calculated a number of statistics to show how the students and NCOs differ in their performance on the new instruments. Descriptive statistics for each of the new instruments are provided in Table 4.16.

These statistics support the notion that the ALQ, LPI, and self-efficacy instruments are measuring something other than cognitive ability and other factors that traditionally relate to effective test performance. We report effect sizes here (rather than the results of significance tests between two means) because effect sizes convey more information about relationships than statistical significance tests do (Schmidt, 1996). The effect size reported in Table 4.16 shows

that undergraduates had lower means and standard deviations for each of these new instruments, especially for Form A of the ALQ and LPI. Therefore, students who might typically perform well in testing situations score lower on the new instruments; their scores also vary less. This finding may be explained by the experience factor: While these students generally lack the knowledge or skills needed to do well on these items, the NCOs possess these characteristics in varying degrees.

Table 4.16
Descriptive Statistics for New Instruments, NCO and Undergraduate Pilot Test Samples

		NCOs			Undergraduates			Effect
Instrument	Split-half r	n	mean	SD	n	mean	SD	Size
ALQ								
Form A	.86	126	40.9	15.8	48	26.1	14.4	.94
Form B	.77	124	36.3	12.8	48	33.9	10.1	.19
Self-efficacy								
Form A	.99	126	747.9	141.7	48	675.4	126.8	.51
Form B	.99	124	736.0	156.7	48	660.9	126.0	.48
LPI								
Form A	.79	115	35.0	15.0	44	16.6	10.6	1.23
Form B	.83	121	28.4	17.2	44	19.3	11.3	.53

Note. Effect sizes were calculated using the formula (XN-XU)/SDN, where N is the NCO sample and U is the undergraduate sample. The larger SD for the NCO sample, rather than the pooled SD, was used in the denominator of the equation.

Correlations among the new instruments and various marker variables for the two samples are presented in Tables 4.17, 4.18, and 4.19. Note that the cognitive marker variables for the NCO sample are verbal, quantitative, and spatial ability test scores, and for the students, GPA, verbal reasoning, and SAT (verbal and math) test scores. The personality/ temperament marker variables were the ABLE subscales for the NCO sample, and social skills subscales for the student sample. An examination of the correlation tables reveals that although the marker variables were different for the two samples, the relationship between the new instruments and the relevant marker instruments was in many cases stronger for the undergraduates than for the NCOs. These statistics supported the notion that for NCOs, the ALQ, LPI, and self-efficacy instruments measured something other than cognitive ability and other factors that traditionally relate to effective test performance. Within the student sample, the ALQ (Form A) and the LPI (Form A) had high correlations with the SAT-Verbal and EAS-VR scores. In addition, these measures showed a strong relationship with the temperament and experience measures taken in the student sample.

Table 4.17

Correlations Between New Instruments and Cognitive Ability Marker Instruments, NCO and Undergraduate Pilot Test Samples

	NCOs			Undergraduates			
Instrument	WK	AR	AO	EAS	GPA	SAT-V	SAT-M
ALQ							
Form A	29	28	20	40	18	55	23
Form B	37	22	26	01	-07	15	-27
Self-efficacy							
Form A	51	27	20	38	-08	53	01
Form B	35	18	18	12	-08	33	09
LPI							
Form A	22	27	01	30	-05	52	30
Form B	20	24	27	27	15	06	-01

Note. Decimal points omitted; WK = Word Knowledge; AR = Arithmetic Reasoning; AO = Assembling Objects; EAS = Logic Test; GPA = College Grade Point Average; SAT-V = Verbal Score on the Scholastic Aptitude Test; SAT-M = Mathematics Score on the Scholastic Aptitude Test.

Table 4.18

Correlations Between New Instruments and Personality Marker Instruments, NCO and Undergraduate Pilot Test Samples

	NCOs									Undergraduates				
Instrument	WO	Adj	Dom	Dep	LC	Cp	PC	UV	NRR	Tol	Dom	SA	SP	OE
ALQ														
Form A	03	03	-02	01	17	06	-08	-13	41	24	39	28	19	28
Form B	25	16	11	17	20	-04	-10	-17	33	01	-14	04	-07	-03
Self-efficacy														
Form A	05	21	20	-23	03	-01	21	04	11	28	40	48	36	32
Form B	12	02	01	-02	12	-14	12	-02	05	15	18	11	13	15
LPI														
Form A	10	08	02	20	12	-10	-06	08	35	32	24	25	13	26
Form B	27	01	06	01	11	-07	-15	-26	55	08	21	17	16	16

Note. Decimal points omitted; WO = Work Orientation; Adj = Adjustment; Dom = Dominance; Dep = Dependability; LC = Locus of Control; Cp = Cooperation; PC = Physical Condition; UV = Unlikely Virtues; NRR = Non-Random Response; Tol = Tolerance for Ambiguity; Dom = Dominance; SA = Social Adroitness; SP = Social Perceptiveness; OE = Oral Expression.

Table 4.19
Correlations Between New Instruments and Selected Background Measures, Undergraduate Pilot Test Sample

Instrument	Age	Sup-W	Sup-C	Mil-Exp	Mil-Fam
ALQ					
Form A	45	56	25	26	31
Form B	42	26	-04	34	25
Self-efficacy					
Form A	31	28	32	27	45
Form B	06	36	-02	13	14
LPI					
Form A	24	31	12	02	24
Form B	25	14	18	12	08

Note. Decimal points omitted; Sup-W = Supervisory Experience at Work; Sup-C = Supervisory Experience Outside of Work; Mil-Exp = Military Experience; Mil-Fam = Familiarity with the Military.
The formula for Hoyt r is: $(MS_B - MS_E)/MS_B$, where MS_B is mean squares between and MS_E is mean squares error.

For the student sample, the moderate to high correlation with verbal reasoning scores makes sense because verbal reasoning should relate to scores on any instrument that requires reading and reasoning. Also, the moderate correlation and with temperament measures such as "dominance" makes sense in that dominance is likely correlated with leadership problem solving skills. Finally, the positive correlations between these measures and supervisory experience is also consistent with our expectations, as the new instruments are basically designed to predict supervisory skill in the Army, and individuals who have been supervisors would likely possess a greater degree of supervisory skill. The pattern of results fits well with what one would expect to find in a sample that lacks the skills purportedly measured by the ALQ and LPI. At a minimum, the results support the notion that the new measures are measuring something other than cognitive ability and temperament in the NCO. Undergraduate students, who tend to lack the skills that the ALQ and LPI purport to measure, should and do obtain scores that correlate highly with the marker variables. In the NCO sample, the correlation between the various marker variables and the new instruments is often lower than that for the undergraduates. These lower correlations are likely due to the fact that the new measures are measuring something other than cognitive ability and temperament in the NCO. The lower correlations cannot be attributed to measurement error in the NCO sample -- the reliability of the marker variables and the new instruments is quite high. Nor can these low correlations be attributed to reduced variance on the new instruments in the NCO sample -- the variance of new instrument scores in the NCO sample is higher than that for the undergraduates.

Therefore, the variance in the NCO scores on the new instruments is likely due to some systematic factor. The fact that NCOs scored higher on these instruments than the

undergraduates lends support to the notion that this factor is related to whether or not one is an NCO. In particular, as NCOs apparently possess a moderate to high amount of tacit knowledge and skills in problem solving, these results suggest that the new instruments measure these constructs—as intended.

CHAPTER 5

REVISE PREDICTOR AND CRITERION MEASURES AND COLLECT VALIDATION DATA

This chapter describes the revisions made to the predictor and criterion measures as the result of pilot testing. Table 5.1 contains a summary of these changes. Then, a brief description of the validation study data collection procedures is provided.

Revisions Made to Predictor Measures

Marker Tests

Verbal Ability

We decided to use the Word Knowledge Test as piloted (20 items) and retain the eight minute time limit, even though all subjects seemed to finish within six minutes.

Quantitative Ability

The decision that we needed to make about the Arithmetic Reasoning Test was whether to shorten the test from 20 items to 14. This is a slightly speeded test; only half of the examinees were able to finish 14 items in 14 minutes, and the hardest items are at the end of the test. The mean number of items completed in 20 minutes was 17.78 (SD = 2.69), and 76% of examinees finished 20 items. Therefore, we retained all 20 items, leaving the test slightly speeded.

Spatial Ability

The Assembling Objects Test was also retained as piloted; examinees must solve 20 problems, ten of each of two distinct types, within a 10 minute time limit.

Biodata Inventory: ABLE

We decided to use the ABLE measure as piloted. Examinees complete the untimed 133 items that comprise nine scales (including a non-random response scale and an unlikely virtues scale).

Background Information Form

This instrument was retained as pilot tested at Fort Hood (this version includes the checklist for types of jobs of the soldiers supervised).

Table 5.1

Summary of Revisions Made to Paper-and-Pencil Instruments Prior to the Validation Study

Measures	Changes Made after Pilot	Time to Complete	#of Items
Background Information Form	Ft. Hood pilot test version selected	5 min. (approx.)	12
Word Knowledge	None	6 min.	20
Arithmetic Reasoning	None	20 min.	20
Assembling Objects	None	10 min.	20
Army Leadership Questionnaire	15 items edited 16 items deleted	60 min. (approx.)	40
Self-Efficacy - Army Leadership Questionnaire	Reduced item total to 20 per form(part of ALQ)		20
Perceived Number of Correct Answers	New	3 min. (approx.)	5
Leadership Problems Inventory	7 items edited 8 items deleted	20 min. (approx.)	133
ABLE	None	25 min.	133
Level of Difficulty Associated with Leadership Problems	New	10 min. (approx.)	5

Note: Instruments are listed in the order they were administered.

New Instruments

Tacit Knowledge - Army Leadership Questionnaire

We made two kinds of changes to items on the pilot version of this instrument--we edited some items and dropped others. We wanted to reduce the number of items on the forms to meet the time objective of a 60-minute test. We determined, from application of the general Spearman-Brown formula (Spearman, 1910; Brown, 1910) that we could reduce the number of items per form to 40 and still maintain a reasonably high internal consistency reliability. Also,

40 items on each form would allow most respondents to complete the form in less than 60 minutes. We went through several steps to arrive at these changes.

First, we reviewed the statistics listed in the results section of this report. These statistics helped us to conclude that:

- tacit knowledge items vary in the amount of agreement between experts on the correct answers;
- the 48 item ALQ has moderately high reliability;
- all scoring methods have essentially the same reliability; and
- scores on the ALQ are related to measures of Non-Random Responding and Cognitive ability.

Second, we reviewed comments received from the USASMA keying sample respondents and recorded them on copies of the A and B forms of the test so that we could review them as a group when we examined the item statistics. We especially took note of items that received multiple and convergent comments. As a result of this review, we edited 15 items (five on Form A and 10 on Form B) and deleted two items (one from each form) that were judged by respondents to be irrelevant problems for NCOs. As examples of editing changes, we:

- changed the word "pass" to "leave" to be consistent in all response options,
- spelled out acronyms correctly, and
- edited out the word "religious" from the phrase "send him to the chaplain for religious counseling."

Third, we reviewed items that had low agreement as to the correct response within the USASMA keying sample. Ten items (four on Form A and six on Form B) had response distributions that were essentially "rectangular" -- among the five response options, no clear choices of "Most Effective" or "Least Effective" could be designated. These 10 items were removed from consideration for the final forms of the ALQ.

Fourth, we further examined the USASMA key and identified two items on Form A that had compound split endorsements (all five response options identified as either a "Most Effective" or a "Least Effective" response, based on the keying criteria). These two items were also deleted from the item pool.

At that stage, we had identified for deletion a set of 14 items from the total set of 96 (48 per original form). Therefore, we still needed to eliminate two items to allow us to balance 40 items between the revised Forms A and B of the ALQ. As a final balancing measure, we swapped several items between the original forms to create more even levels of item difficulty across the eight item categories and moved several items from one form to the other to maintain equivalency of item categories between the forms. In conducting this step, two items were deleted to increase the degree of balance.

Finally, we ordered the items in terms of difficulty on both forms. To do this, we used scoring Algorithm 3, as this method was the most straight-forward of the four, and the reliability of the scoring alternatives remained essentially the same.

Self-Efficacy - Army Leadership Questionnaire

We changed our approach to measuring this variable in several ways. In particular, we:

- retained our original measure of self-efficacy, but reduced the number of ALQ items for which it would be collected.
- developed a new measure called "Perceived Number of Correct Answers-- Army Leadership Questionnaire" meant to reflect self-efficacy on the ALQ in general. D
- developed a new measure called "Level of Difficulty Associated with Leadership Problems" meant to examine self-efficacy with handling problems on the job as an NCO.

Self-Efficacy - Portion of ALQ. In reviewing the pilot data from the self-efficacy instrument we found that the variance between respondents was high, the variance within respondents was low, and the reliability of the responses was high.

Therefore, we concluded that it was possible to reduce the amount of time that respondents spent answering this type of item, and still have a highly reliable self-efficacy measure. Reducing the time spent on this measure allowed us to obtain additional measures of self-efficacy.

We applied the general Spearman-Brown prophecy formula (Spearman, 1910; Brown, 1910) and found that the self-efficacy scales would maintain reliability in the mid-90's even after reducing the number of items to 20. Therefore, we decided to reduce the number of ALQ items for which the measure is applied to 10 (providing 20 self-efficacy items--2 for each problem scenario). This allowed us to cut the amount of time spent on the measure in half, continue to address a wide range of scenarios that NCOs typically face, and maintain a high internal consistency reliability for the measure. We chose the 10 problems so that they varied in both item difficulty and problem dimension. The ALQ items on each form were ordered so that these 10 problems were presented first.

Perceived Number of Correct Answers. We developed this measure in an attempt to measure self-efficacy on the ALQ through a one-page questionnaire that would follow the ALQ. The measure asked respondents to indicate how confident they were that they gave a certain percentage of correct answers on the ALQ.

Note that this measure:

- like the original self-efficacy measure, pertained to self-efficacy on the ALQ instrument
- treated the ALQ as a task, while the original self-efficacy measure treats each item as a task. This is important because researchers (e.g., Bandura, 1986) have recommended that self-efficacy be examined on a task-specific basis.
- used the "self-efficacy composite" score method recommended in the self-efficacy literature (Lee & Bobko, 1994).

Level of Difficulty Associated with Leadership Problems (LDLP). Critics of the first two measures of self-efficacy may claim that these measures should not be used to predict NCO performance because they do not involve judgments of on-the-job performance. However, the tasks depicted on these instruments were carefully developed to represent decisions often required of NCOs, and decision making is a major part of any NCO's job. In addition, in this potentially high-stakes assessment context, it may be difficult to obtain useful data through a direct measure of on-the-job self-efficacy. That is, if NCOs were asked about how they think they would perform as an NCO, it is likely that scores would be inflated and demonstrate little or no interpretable variance. Still, we decided it would be prudent to develop a self-efficacy measure that pertained to on-the-job NCO performance, in case there was sufficient time to administer it.

To develop this measure, we pulled five problem scenarios from Form B of the ALQ that varied in terms of difficulty and problem dimension. Then, we presented these problems without potential solutions, and asked examinees to respond to a series of questions about their confidence in dealing with problems like these. The response format was similar to the one recommended in the literature (Lee & Bobko, 1994), in which a respondent provided two responses for several levels of probability of achieving a goal (e.g., solving a problem):

- As an NCO, how often could you deal effectively with a problem like this? (checked Yes or No for each of five probability levels -- ranging from 50% of the time to all of the time)
- How sure are you? (rated on a 9 point scale: 1= Not Sure; 9=Very Sure)

To deal with potential problems with "faking good" on this type of measure, we asked respondents to focus on specific tasks. Our rationale was that the details of each scenario may lead individuals to respond honestly to the confidence judgments.

Priority Setting: Leadership Problems Inventory

We made two kinds of changes to items on the pilot version of this test -- we edited some items and dropped others. We went through a series of several steps to arrive at these changes.

First, we reviewed the statistics listed in the results section of this report. These statistics helped us to conclude that:

- LPI items varied in the amount of agreement between experts on the correct answers.
- the 24 item LPI had moderately high reliability.
- some scoring methods were more reliable than others.
- scores on the LPI were related to measures of Non-Random Responding and Cognitive ability.

Then, we applied the general Spearman-Brown prophecy formula (Spearman, 1910; Brown, 1910) to determine the number of items that we could drop on the LPI and still maintain a reasonably high reliability. We estimated that we could reduce the number of items to 20 on each form and maintain a reliability greater than .80.

Next, we reviewed comments received from the USASMA keying sample and recorded them on copies of the A and B forms of the test, so that we could review them as a group when we examined the item statistics. As a result of these comments we made minor wording changes to one item, and we decided to search for items that were problematic because field and garrison problems were mixed. The NCOs found it difficult to consider the priority of problems in a block simultaneously when some of the problems occurred in completely different contexts. This was particularly a problem when some of the problems in a block occur in garrison, and some of them occur in the field.

Therefore, we reviewed each problem used in the LPI and determined if the problem would occur in the field, garrison or some other specific location. As a result of this review, we found five problems that would occur in the field, while the remainder could occur anywhere or at garrison. These five field problems were used in 12 separate LPI items. These items were marked for editing or deletion.

Then, we reviewed statistics comparing the December USASMA LPI priority rankings to the LPI priority rankings based on priority scores obtained at the May, 1994 USASMA data collection. We identified items with poor agreement between these two expert samples, and marked them for possible deletion.

Next, we selected items for the final two forms (16 from Form A, 17 from Form B) that had reasonable agreement between the two expert samples, varied in terms of distribution of priority scores, and did not have field problems included.

At this stage, we had a total of 33 items (16 from Form A, 17 from Form B) to distribute to two LPI forms. We needed seven additional items to have the desired total of 40 to distribute equally among the two forms. However, each of the seven items we wished to add had only four problems to prioritize, instead of the five problems appearing in all the other 33 items. We therefore completed each of these seven four-problem sets with a garrison problem, so that all 40 items would have five problems, and all problems in the seven items would be garrison problems. Seven specific garrison problems were selected as replacements that: (a) had a similar priority score and (b) were from the same problem category as the problem they replaced. We decided to exclude subjects' responses to these replacement problems when scoring the seven items, since we did not have keying data available for the newly-created problems sets.

Thus, these seven problems are scored as if only the four keyed problems were ranked by subjects.

We decided that Form B should be used as the primary form in the validation study as there were fewer changes to Form B following pilot testing than to Form A. Finally, the items were rearranged within forms so that they could be presented in descending order of agreement between the two expert samples. This is based on the reasoning that those items that have a higher level of agreement between experts would be less difficult.

Revisions Made to Criterion Measures

The following is a summary of the revisions made to criterion measures based on pilot test results. These final versions of the criterion measures were used in the validation study.

Personnel File Form

Examination of the item statistics from the pilot indicated that all items had reasonable response distributions, suggesting that no changes to the form were necessary. However, post-pilot test interviews with NCOs indicated that there were a few points of confusion to remedy. In light of these comments, we made three significant changes to the PFF prior to beginning the validation study data collection.

The first significant change was to modify the response format for the Awards/Badges/Commendations section (item #1) to require respondents to answer either "yes" or "no" to each listed award, badge, and commendation. This change allowed us to determine (during the data collection sessions) that respondents had completed the section; with the previous format, we could not determine (with complete certainty) the difference between missing data and lack of awards received.

The second significant change was to replace the response options for items 1q, 1r, and 1s (asking about the number of various medals received) from "1", "2", or "3 or more" to "0", "1", "2", "3", "4", or "5 or more", and to require all respondents (even those who had never received a given medal) to respond. The "0" response was added to be consistent with later items in which the respondent was required to make a frequency response, even when the answer was 0; the "4" and "5 or more" options were added to eliminate a potential ceiling effect indicated by the pilot test data.

The last significant change was to replace items 4a, 4b, and 4c (asking about numbers of business school, trade school, and college courses completed) with a seven-item checklist (4a-4g) of military courses completed. This was the same checklist that had been used in the Project A/Career Force version of the PFF.

Other minor changes were made to some of the items to make the wording more similar to that used in Project A/Career Force. For instance, the phrase "Check (✔) the awards and

decorations listed below that you have received" was changed to "Put a check (✔) next to the awards and decorations listed below that you have received."

NCO Behaviorally Anchored Rating Scales

Two groups of six participants reviewed the NCO BARS on the second and third days of the pilot test at Fort Hood. In two cases suggestions led to the revision of incidents that were included as anchors for the low, moderate, and high ratings on performance dimensions. Comments were also used to modify the wording on the dimensions to resemble more closely the responsibilities of the NCO on the job. Following the pilot test and discussion, revisions were made to the NCO BARS according to suggestions from the participants.

Structured Interview

Interviews were conducted over three days. Half of the interviews were videotaped to develop interviewer training tapes for later use. Generalizability analyses were conducted to determine reliability of the interview questions. Single rater reliability was .77 for making relative decisions (decisions between NCOs across the 5 dimensions). This was considered adequate reliability for the Structured Interview measure.

Minor revisions to the wording and the order of the interview questions and the scales were made and tested throughout the pilot test. For instance, it was determined that to make the interview questions easier to answer, the interview should begin with questions about motivating others. Most NCOs had experience on this dimension, which made these questions easier to answer and, therefore, started the interview off on a positive note.

Job Analysis Questionnaire

This instrument was retained as piloted.

Collect Validation Data

The purpose of the validation data collection was to collect the data necessary to evaluate the validity of the new measures for predicting NCO performance, and the degree to which the new instruments increment the prediction of NCO performance over already available or traditional measures. The validation data collection took place at five sites during fiscal years 1995 and 1996. Our goal was to obtain complete data at each site from 150 NCOs representing a variety of MOS and the ranks of E-5 through E-8.

Data collection at each site required five days to complete, and consisted of three basic activities: paper-and-pencil testing, structured interview, and collection of supervisory ratings. Paper-and-pencil instruments were administered in the order specified in Table 5.1. At the time of paper-and-pencil testing, an appointment for the structured interview was made for a one hour time interval later in the week. Supervisors of NCO participants were also contacted at that time so that appointments could be made for them to provide ratings of the NCO participants. At the

last site, we also obtained a second supervisory rating for a portion of the sample, to use to calculate an interrater reliability coefficient.

Prepare for Validation Data Collection

To prepare for collecting the validation study data, we contacted the sites to arrange for facilities and participants. We also prepared materials and protocols for the testing sessions, structured interviews, and supervisor rating sessions. An essential part of preparing for the data collection was developing scripts for administering the three types of sessions. Finally, we conducted a training session to prepare the data collection team for the on-site activities.

We provided our Points of Contact (POCs) at the five sites with a manual explaining the data collection requirements, methods, and procedures. We made arrangements with our Points of Contact (POCs) at the sites to provide personnel resources and facilities as follows:

150 NCO Validation Study Participants

- 100 NCOs at the E-5/E-6 level
- 50 NCOs at the E-7/E-8 level
- Approximately one-third of the 150 total representing each of the three general MOS groupings: (1) Combat, (2) Electronic/Repair, and (3) Administrative/Support

NCO Supervisors

- One direct supervisor for each of the 150 E-5 through E-8 level NCOs (Note that some supervisors ended up rating more than one of the participating NCOs. Some NCOs included in the validation study also actually supervised other study participants, and therefore also provided ratings of their subordinates.)

Facilities

- One large classroom suitable for seating 75 NCOs at a time, for two four-hour written testing sessions
- One medium-sized room suitable for seating up to 20 supervisors at a time
- Six small, office-sized rooms, each with two chairs and a desk

Prepare Materials and Script

We prepared a testing order for the written measures session and copies of all instruments and answer sheets for collecting written predictor and criterion measures. We also prepared necessary procedures and materials for the supervisory rating sessions and structured interviews, including supplies for videotaping approximately two interviews per hour.

In preparation for the training sessions we held for staff, we developed scripts for the three segments of the data collection: the written measures sessions, the supervisor ratings sessions, and the structured interview. These scripts included a brief overview of the project

background and purpose and a description of the specific steps involved in completing the specific activity.

Conduct Training for Data Collection Team

We held a one-and-a-half day training meeting to prepare for the data collection trips. The objectives of the meeting were to familiarize all staff with the general and specific procedures for conducting the written and interview sessions, to take turns practicing the interview questions and probes, and to check the consistency with which we used probes and rated interview responses. We spent approximately one half day covering the written measures administration and devoted a full day to training for the structured interviews, to ensure consistency among staff members in conducting the interviews and scoring responses.

Training for Written Measures Sessions

The training for the written sessions was straightforward. We reviewed the project purpose and general testing issues (e.g., test security, handling privacy and confidentiality concerns). We explained what each instrument measured and whether it was timed vs. untimed, described the test order, and reviewed the instructions for each test. We also reviewed the procedures developed for scheduling interviews and verifying supervisor information.

Training for Structured Interviews

Prior to the training session, we prepared a "Structured Interview Administration and Assessment Procedures" manual containing all necessary procedures and materials for conducting the ECQUIP structured interview. The manual described the major steps in the interview process:

- preparing for the interview (setting interview times, organizing the interview context)
- opening the interview (establishing rapport, explaining the interview structure)
- gathering information (probing to understand the interviewee's responses, taking notes, following instructions for questions that tap multiple dimensions, closing the interview)
- evaluating the NCO (making ratings, avoiding common rater errors)

During the training session, we reviewed as a group the purpose and structure of the interview and then systematically reviewed all sections of the manual. This included covering specific instructions for probing and taking notes, explaining common interviewing problems, and completing exercises to phrase probes and judge responses. The group also reviewed the individual rating scale definitions and anchors to come to a common understanding of the types of responses that qualify at the low, moderate, and high points of the scales. We also devoted time to describing procedures specific to conducting videotape interviews, e.g., obtaining consent from an NCO before starting the interview, handling questions about how the tapes will be used and who will view them, etc.

Training for Supervisor Rating Sessions

This training was not conducted as part of the session because only one person on each data collection trip was responsible for gathering supervisor ratings. Therefore, only two members of the data collection team learned these procedures, with the intent of alternating the role for the total of five data collection trips.

Data Collection Procedures

We collected data at three sites during fiscal year 1995 and two additional sites during fiscal year 1996. At each site, we followed the schedule and procedures described below. In some instances, we deviated from these procedures to accommodate the unique needs and constraints at a site.

Written Sessions

Written measures were administered in two four-hour blocks the first two mornings at each site. Half of the NCOs attended each session. Each written session began with a description of the project background and the NCOs' role in the project. We passed out the Privacy Act statement and had NCOs complete the Background Information Form.

We first conducted the timed tests (verbal, quantitative, and spatial) and then passed out an envelope containing the remaining (untimed) materials to be completed during the session. As a group, we reviewed the instructions for the first untimed measure, the Army Leadership Questionnaire, and then asked the NCOs to continue working individually. The order for the timed and untimed measures is provided in Table 5.2.

Table 5.2

Administration Order for Written Measures Session

Measures	Time to Complete	# of Items
Background Information Form	5 min. (approx.)	12
Word Knowledge	6 min	20
Arithmetic Reasoning	20 min.	20
Assembling Objects	10 min.	20
Army Leadership Questionnaire	60 min. (approx.)	40
Self-Efficacy - Army Leadership Questionnaire (part of ALQ)		20
Perceived Number of Correct Answers	3 min. (approx.)	5
Leadership Problems Inventory	20 min. (approx.)	133
Personnel File Form	5 min. (approx.)	11
JAQ	10 min. (approx.)	13
ABLE	25 min.	133
Level of Difficulty Associated with Leadership Problems	10 min. (approx.)	5

While NCOs completed the untimed portion of the written measures, we circulated around the room and scheduled each NCO for an appointment to return to complete the structured interview. At this time, we also verified that we had the correct supervisor name and phone number for each NCO.

Structured Interviews

When an interviewee arrived at the scheduled interview time, an available interviewer introduced him/herself and selected an open interview station. If one of our two video stations was not in use, the interviewer started the interview at that station, since a primary objective was to collect as many videotaped interviews as possible with the two cameras. If both of the video stations were in use at the time, the interviewer situated the interview at one of the other (no camera) stations.

Each interview started with an explanation of the purpose of the interview. For those NCOs at a station with a camera, the purpose of videotaping was also explained. To break the ice and get the NCO talking, the interviewer asked the NCO to briefly describe his/her current job. This brief description also allowed the interviewers to establish some context for the answers the NCO would give to the nine interview questions.

Several specific instructions were read verbatim from the script, including the desired structure of (1) situation, (2) action, and (3) result for each response. The interviewer then gave a very simple example of both a question and an appropriate answer, using the specified structure:

Question:	*Suppose I was interested in your driving experiences. I might ask you a question like: tell me about a time when you avoided a dangerous situation while driving. Your response might be:*
Answer:	*Situation: I was driving at night and it was raining when another car pulled out in front of me.* *Action: I anticipated the other driver's actions, so I put on my brakes and swerved to avoid the car.* *Result: An accident was narrowly avoided.*

After the instructions were read, the actual interview started; for NCOs at video stations, the camera was turned on at this point. Each interviewee was asked to answer nine questions that tap the four dimensions (1) Motivating Others, (2) Demonstrating Effort, (3) Planning and Providing for Training, and (4) Organizing, Coordinating, and Executing. The interview questions for each dimension are listed below.

- **Motivating Others**

1. *I'd like you to tell me about a time when you thought something was wrong with a soldier. What made you think there was a problem?*

2. *Now, I'd like you to tell me about a time when you had to criticize a subordinate. What exactly did you say?*
3. *Describe a time when you had to encourage one or more of your subordinates to perform an unpleasant task. What did you do?*

- **Demonstrating Effort**

4. *Tell me about a time when you put forth a great deal of effort to accomplish a task or activity at work.*

5. *Now, tell me about a time when you had to take the initiative to get something important done.*

- **Planning and Providing for Training**

6. *Tell me about a time recently when you had to train some subordinates. Describe the steps you went through to conduct the training.*

7. *Tell me about a time when you were training soldiers to do something that was difficult or complicated to master. What did you do?*

- **Organizing, Coordinating, and Executing**

8. *Now, tell me about a time when you encountered a problem at work. What alternatives did you consider to solve it?*

9. *Describe an activity where you coordinated people, equipment, or some other resource.*

For each question, the NCO was asked to first take a minute to think of an appropriate situation and then provide the three-part answer. Interviewers used prompts and probes as necessary to elicit information to take notes on the situation, action, and result. If an NCO was unable to think of an example situation, he/she was given the option of continuing with the other questions and returning to the earlier question(s) later. After completing all nine questions, the interviewer thanked the NCO and stopped the camera, if applicable.

Immediately after each interview, the interviewer reviewed his/her notes and referred to the behavioral anchors on the performance scales to evaluate the behavioral responses. Each NCO was rated on four dimensions which correspond with the four main question areas listed above. In addition, each NCO was rated on a fifth dimension of Communicating Orally. This provided an evaluation of the NCO's professionalism and ability to express him/herself in a logical, organized, and confident manner.

Supervisor Rating Sessions

As soon as the verification was completed, we began the process of collecting supervisor ratings. The basic procedures for collecting supervisor ratings data were: ensure the accuracy of the supervisor name for each NCO, prepare a rating packet (with all applicable supervisees) for each supervisor, schedule supervisors to attend small group sessions, conduct rating sessions, and track and contact missing supervisors.

For each session, the raters were given a brief explanation of the rating process and how the data were to be used. They received an instruction packet, a set of rating scales, and a sheet for recording ratings. Each individual received a card which listed the names of the NCOs to be rated; this card matched up with the lines on the rating sheet to ensure that the process for recording the ratings was accurate and easy. Supervisors were asked to work individually to complete the ratings, and to ask for help if necessary. For each NCO in the study, we collected supervisor ratings for thirteen performance dimensions plus eight supervisory problem categories.

Data Collection Sites and Numbers

We tested at the following sites during their scheduled umbrella weeks:

Site	Date
Fort Eustis	March 20 - 24, 1995
Fort Bragg	March 27 - 31, 1995
Fort Campbell	June 5 - 9, 1995
Fort Drum	January 16 - 19, 1996
Fort Hood	March 18 - 22, 1996

At the end of the 1995 fiscal year, we evaluated our actual sample against the original planned sample. As expected, administrative/support MOS were under represented, reflecting the actual distribution of MOS in the Army. An over representation of electronic/repair MOS after collecting data at the first three validation sites was due to our inclusion of Fort Eustis, which had a high percentage of personnel in non-combat MOS. Therefore, we requested that the remaining two sites provide proportionally more NCOs from the administrative/support MOS, and targeted the two sites that had a high representation of combat MOS. Table 5.3 shows the representation of MOS groupings that we achieved by the end of the data collection phase.

Table 5.3

NCO Participants by MOS Grouping

MOS Grouping	Sample Size -- Goal	Collected Sample Size
Combat MOS	250	267
Electronic/Repair MOS	250	231
Administrative/Support MOS	250	172
Total	750	670*

* The total number of NCOs who reported MOS grouping on the background information form was 670.

CHAPTER 6

VALIDATION ANALYSES

Chapter 5 described the collection of the validation study data. In this chapter we describe and discuss the methods and results of analyses of those data. The methods and results are presented and discussed in five major sections, after a brief description of the sample. The first two sections describe the predictor and criterion analyses, respectively. The predictor analyses include descriptive statistics, reliabilities as appropriate, and correlations between predictor measures. The criterion analyses include the same types of analyses, but also describe the formation of a single criterion composite score for use in validation analyses.

The third major section describes the analyses of the relationships between the predictor and criterion measures--the validities. These analyses include:

- simple, uncorrected bivariate correlations between the market and new predictors and the criteria measures,
- multiple regressions corrected for range restriction and adjusted for shrinkage, used to estimate incremental validities of each new predictor, and
- multiple regressions, corrected and adjusted, used to estimate the absolute and incremental validity of the "optimum" or best set of new predictors.

The fourth section describes the investigation of validity generalization across job-related subgroups. It includes cluster analyses of the sample of soldiers into three sub-groups using the Job Analyses Questionnaire and a meta-analysis of validity across these three subgroups. The last section describes fairness analyses of selected predictor composites that include the AFQT and the most promising of the new predictors.

Sample

A total of 691 NCOs participated in the study. The sample consisted of 35 females and 633 males, with 23 missing data for gender. The mean tenure in the Army was 11 years and 11 months. The breakout by site is reported in Table 6.1. The breakout by race is reported in Table 6.2. Breakout by major job type is shown in Table 5.3 in Chapter 5.

The ranks for study participants are reported in Table 6.3. The ranks of the participants in this sample are somewhat higher than those for the NCO pilot test sample (see Table 4.3 in Chapter 4), with roughly twice the proportion of E-7's and E-8's.

Table 6.1
Site Frequency and Percent of Sample.

Race	Frequency	Percent
Fort Bragg	126	18.9
Fort Campbell	105	15.7
Fort Drum	152	22.8
Fort Eustis	147	22.0
Fort Hood	137	20.5

Note: Site data were missing for remainder of sample.

Table 6.2
Race Frequency and Percent of Sample.

Race	Frequency	Percent
White	345	51.7
Native Am./Am. Indian	12	1.8
Hispanic	80	12.0
African-American	201	30.1
Asian	15	2.2
Other	14	2.1

Note: Race data were missing for remainder of sample.

Table 6.3
Rank Frequency and Percent of Sample

Rank	Frequency	Percent
E-5	304	45.7
E-6	172	25.9
E-7	164	24.7
E-8	25	3.8

Note: Rank data were missing for remainder of sample.

Predictor Analyses

Data Screening

Our first step in conducting the predictor analyses was to screen the data. First, if a respondent had not provided a response to any of the items on an instrument, then we set that predictor instrument score to missing. With the exception of the ABLE scale scores, this was the only rule used to set predictor instrument scores to missing. All ABLE scale scores were set to missing if a) the percentage of item responses missing on the entire instrument was greater than 10 percent, or b) the number of correct answers to the non-random response scale was less than 75 percent. In addition, an ABLE sub-scale score was set to missing if the percentage of item responses missing on the sub-scale was greater than 14 percent.

ASVAB Scores

One of the major goals of this research was to determine the degree to which the experimental predictors predict performance above and beyond existing operational measures. For this reason, we attempted to obtain ASVAB and AFQT scores for every NCO in the sample. Factor analyses of the ASVAB subtest correlations typically result in four factors (Kass, Mitonell, Gaggerr, & Wing, 1983). We computed four composites of ASVAB scores to reflect the four factors of "factor scores." Table 6.4 shows the descriptive statistics for the ASVAB factor scores and AFQT scores for the NCOs in our sample. We found that ASVAB scores were available for over 86 percent of the sample, however, not all of these scores could be standardized to the same metric. ASVAB scores that could be standardized to the 1980 Youth Population (DoD, 1982) were available for only 66 percent of the sample. The remainder of the sample for whom we were able to obtain ASVAB scores had scores that could be standardized to the 1944 norms only. The descriptive statistics for ASVAB factor scores shown in Table 6.4 are for that part of the sample where ASVAB scores could be standardized to the 1980 Youth Population. We were able to obtain AFQT scores standardized to the 1980 population for nearly 90 percent of the sample.

Marker Tests

Descriptive statistics and reliability coefficients for the Word Knowledge Test, the Arithmetic Reasoning Test, the Assembling Objects test, and the ABLE are provided in Table 6.5. In general, these statistics are similar to those found for these instruments in the NCO pilot test, (see Table 4.4) although the cognitive test scores are somewhat higher in this sample. This difference is likely due to the fact that there are proportionally more high ranking NCOs in this sample than in the pilot sample.

Table 6.4

ASVAB and AFQT Descriptive Statistics.

	n	Mean	SD
ASVAB Factor: Technical	458	135.99	17.94
ASVAB Factor: Quantitative	458	104.45	13.87
ASVAB Factor: Verbal	458	104.06	13.13
ASVAB Factor: Speed	458	107.28	11.69
AFQT Score	592	52.16	22.84

Table 6.5

Descriptive Statistics for Marker Tests.

Marker Test	n	Mean	SD	Hoyt r
Word Knowledge	668	14.69	2.99	.71
Arithmetic Reasoning	670	12.96	3.69	.78
Assembling Objects	669	14.78	3.67	.80
ABLE				
Work Orientation	629	68.91	7.58	.87
Adjustment	629	35.92	4.77	.80
Dominance	629	46.41	5.15	.80
Dependability	628	51.99	6.00	.81
Locus of Control	629	31.66	3.71	.68
Cooperation	629	21.54	2.24	.37
Physical Condition	629	18.47	3.60	.83
Unlikely Virtues	629	16.04	3.06	.63
Non-random Response	667	7.54	1.05	.69

Predictors from the Personnel File Form (PFF)

Some items from the Personnel File Form reflect factors that likely contribute to the selection of NCOs. The items that may tap these factors are listed in Figure 6.1. Each of these items consists of a count of the number of letters, disciplinary actions, etc. at each rank. We examined these items and developed a model as to how they might co-vary. We hypothesized that items 2a and 3a would load on a factor that we might call "Recognition", and that items 10a and 11a would load on another factor that we might call "Disciplinary Actions". Finally, we hypothesized that item 7 (through Grade E-4 only) could form its own factor called "Recommendations for Accelerated Promotions". We tested this model in a confirmatory factor analysis, and found that the model indeed fit ($\chi^2 = 6.53$, Root Mean Squared Residual = .03). Thus we formed three additional nonexperimental predictors by summing the items that load on each factor.

Figure 6.1
PFF items used.

2. Check (✔) the number of Memoranda/Letters of Appreciation, Commendation, or Achievement you have received....

   a. while in grades E-1 to E-4: __0 __1 __2 __3 __4 __5 __6 __7 or more
   b. while in grades E-5 or above: __0 __1 __2 __3 __4 __5 __6 __7 or more

3. Check (✔) the number of Certificates of Appreciation, Commendation, or Achievement you have received....

   a. while in grades E-1 to E-4: __0 __1 __2 __3 __4 __5 __6 __7 or more
   b. while in grades E-5 or above: __0 __1 __2 __3 __4 __5 __6 __7 or more

7. Have you ever been recommended for an accelerated promotion (i.e., promotion in the secondary zone)?

   ___Yes ___No

   If yes, check (✔) the paygrade(s) that you were in when the recommendation(s) were made. Check all that apply.

   ___E1 ___E2 ___E3 ___E4 ___E5 ___E6 ___E7 ___E8

10. Check (✔) the number of Articles 15 you have received....

    a. while in grades E-1 to E-4: __0 __1 __2 __3 __4 or more
    b. while in grades E-5 or above: __0 __1 __2 __3 __4 or more

11. Check (✔) the number of Flag Actions (i.e., suspension of favorable personnel action) you have received....

    a. while in grades E-1 to E-4: __0 __1 __2 __3 __4 or more
    b. while in grades E-5 or above: __0 __1 __2 __3 __4 or more

Descriptive statistics for the three measures derived from the PFF are shown in Table 6.6. Note that Disciplinary Actions are highly skewed, so that more than 70 percent of the sample has a score of 0. However, as Disciplinary Actions are particularly serious in the Army, we chose to retain the measure despite its relatively poor distributional properties.

Table 6.6
Predictor Measures From Personnel File Form: Descriptive Statistics.

	n	Mean	SD
Recognition	644	7.11	3.83
Disciplinary Actions	644	0.32	0.83
Rec for Accel Promo	618	0.63	0.48

Correlations among Nonexperimental Predictors

Table 6.7 presents the correlations among the nonexperimental predictors. These correlations fit the expected pattern in that scores meant to reflect the same or similar constructs correlate highly. That is, we obtained high correlations among measures of g, and particularly high correlations among measures of verbal ability, and among measures of quantitative ability.

The high correlations between the ASVAB scores and the cognitive marker tests (WK, AR, AO) show that there has been little decay in the relevance of the ASVAB scores in this sample, despite the fact that most of these scores date back to the soldiers' entrance into the Army, on average, 11 years ago.

The pattern of correlations between the cognitive test scores and the ABLE measures are similar to those found in the concurrent validation study conducted as part of Project A (McHenry, Hough, Toquom, Henson & Ashworth, 1990). Finally, the PFF measures of Accelerated Promotion showed considerate positive relationships to most of the cognitive test scores. The slightly negative relation between Recognition and Cognitive test scores is puzzling but not problematic.

Experimental Measures

Recall that we developed five experimental measures: the Army Leadership Questionnaire (ALQ) (Forms A and B), the Leadership Problems Inventory (LPI) (Forms A and B), the Self-efficacy measure from the ALQ (SE-ALQ) (Forms A and B), the Perceived Number of Correct on the ALQ (PNC), and the Level of Difficulty Associated with Leadership Problems (LDLP). First, we developed and pilot tested two forms each of the ALQ, LPI, and SE-ALQ. Then we administered these forms to our keying sample. Then, informed by the pilot test and keying sample results, we created two revised forms of each of these measures. In revising the forms, we eliminated items that had poor agreement among experts, and we balanced the items on each form to ensure that each form had similar items in terms of content and difficulty. Next, we developed the PNC and LDLP to serve as additional measures of self-efficacy. The descriptive statistics and reliability coefficients for the experimental measures in the validation sample are provided in Table 6.8.

Note that we report statistics only for Form A of the ALQ, LPI, and SE-ALQ as these were the forms administered to the validation sample, except for a small subsample at the one site that received both forms. As the PNC and LDLP were developed following the pilot test, scoring issues on these instruments had to be resolved on the basis of the validity study data. We will discuss the various scoring methods that we considered, as well as our rationale for our choice later in this section.

Tacit Knowledge: Army Leadership Questionnaire.

Following the recommendations from the pilot test, we analyzed ALQ data using ALQ Scoring Algorithm 4. The mean ALQ score is higher in this sample than in the pilot test despite the fact that there are eight fewer items in this version. Also, the reliability of the ALQ is lower _

Table 6.7

Correlations Among Nonexperimental Predictor Measures.

Predictor	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
ASVAB																			
1. AFQT	--																		
2. Technical Factor	63	--																	
3. Quantitative Factor	89	50	--																
4. Verbal Factor	84	69	59	--															
5. Speed	18	-01	23	03	--														
Other Cognitive Markers																			
6. Word Knowledge	63	52	45	71	01	--													
7. Arithmetic Reasoning	64	47	67	51	15	41	--												
8. Assembling Objects	43	51	39	35	11	30	45	--											
9. Work Orientation	-04	03	-01	-01	05	06	04	02	--										
10. Adjustment	08	06	12	06	05	02	09	00	25	--									
11. Dominance	02	02	00	03	04	07	02	-02	61	29	--								
12. Dependability	-03	-09	08	-09	07	-04	02	-09	25	24	14	--							
13. Locus of Control	05	09	06	05	-01	01	09	04	38	28	24	30	--						
14. Cooperation	-03	-06	-07	-04	11	-04	-02	-00	27	-09	23	-00	11	--					
15. Physical Condition	-07	-11	-08	-11	02	-07	-08	-15	23	17	38	03	-00	11	--				
16. Unlikely Virtues	-29	-21	-17	-30	00	-27	-15	-11	34	16	23	31	14	05	12	--			
17. Non-Random Resp	18	21	16	20	09	19	20	21	01	05	07	07	07	-01	-02	-13	--		
Personnel File Form																			
18. Recognition	-10	-05	-15	-08	01	-02	-07	-07	14	01	21	04	04	02	11	10	-07	--	
19. Disciplinary Actions	-02	-02	-04	-00	-02	-00	-00	-00	-02	04	-01	-15	-03	03	-00	-08	00	02	--
20. Rec for Accel Promo	26	10	18	25	04	26	19	12	18	00	24	03	-01	07	07	-08	02	01	-08

Note. Decimal points omitted; "--" = 1.0; n ranges from 430 to 670.

in this sample (the reliability estimate of the pilot test version of the ALQ, adjusted to a length of 40 items, was .82).

Table 6.8

Descriptive Statistics and Reliability Estimates for Experimental Predictors.

Predictor	N	Mean	SD	Hoyt r
ALQ (Form A)	653	42.40	10.11	.71
LPI (Form A)	669	30.07	10.14	.71
SE-ALQ (Form A)	626	161.62	34.55	.96
PNC	641	2.98	1.19	.67
LDLP	638	3.80	1.17	.86

One explanation for the lower reliability of the validity study version of the ALQ is the manner in which we selected items from the pilot test version. Recall that we deleted items that had low agreement among the USASMA keying study participants. While this is a reasonable strategy, it resulted in the elimination of many items with low mean values in the pilot test sample, as these items were apparently more ambiguous to the NCOs in the pilot test as well. Thus, the sample mean and variance on this version of the ALQ is lower, and the reliability estimates are lower.

In order to examine the psychometric characteristics of Form B of the ALQ, a small subset of the validation study participants ($n = 57$) was made available to us. Descriptive statistics for Form B were comparable ($M = 43.0$, $SD = 7.9$) to those for the same study participants on Form A ($M = 45.79$, $SD = 9.17$) and, although the reliability estimates for Form B were low (Hoyt $r_{xx} = .54$, Split-Half $r_{xx} = .37$), the correlation between the two forms was reasonably high ($r = .64$). Conclusions that can be made regarding Form B are limited because of circumstances unique to its administration. Form B was administered at only one site, on a Friday afternoon, after all other data collection activities were completed. The soldiers seemed to resent having to take another form of a test they had taken earlier in the week. Many of them appeared to hurry through the task. For this reason, we recommend further evaluation of the ALQ-Form B with other samples.

Leadership Problems Inventory (LPI).

Following the recommendations from the pilot test, we analyzed LPI data using LPI Scoring Algorithm 4. The mean LPI score is in line with that from the pilot test data, considering the fact that there are four fewer items in this version. However, the standard deviation of the responses and the reliability estimate are somewhat lower than what we expected (the Hoyt (1941) reliability estimate of the pilot test versions of the LPI, adjusted to a length of 20 items, was .75 for Form A and .80 for Form B). As with the ALQ, the slightly reduced variance and lowered reliabilities in the validation study version of the LPI is likely due to deletion of the more "difficult" items when we constructed the validity study version of the LPI.

We also administered Form B of the LPI to a subset of the validation study participants (n = 63). Descriptive statistics for Form B were comparable (M = 34.02, SD = 5.21) to those for the same study participants on Form A (M = 32.52, SD = 8.68) and the reliability estimates for Form B were again poor (Hoyt r_{xx} = .18, Split-Half r_{xx} = .21). The correlation between the two forms was moderate (r = .42). Generalizations from these results are limited, however, as the LPI-Form B was administered under the same circumstances as the ALQ-Form B.

Self-Efficacy: Army Leadership Questionnaire.

Following the recommendations from the pilot test, we computed a total SE-ALQ score by summing the confidence ratings for the least response and the most response for all 10 items that included the confidence ratings. This still provided us with a highly reliable measure, as expected.

We also administered Form B of the ALQ to a subset of the validation study participants (n = 48). Descriptive statistics for Form B (M = 178.21, SD = 27.08) show that there is less variance in the ratings than those provided by the same study participants on Form A (M = 164.85, SD = 42.62). Also, although the reliability estimates for the Form B were high (Hoyt r_{xx} = .98), the correlation between the two forms was only moderately high (r = .55). Again, however, the Form B results are limited by the poor circumstances under which the test was administered.

Perceived Number of Correct (PNC)

As discussed in Chapter 5, we developed two additional measures of self-efficacy following the pilot test: PNC and LDLP. These instruments were additional attempts at tapping self-efficacy. For each of these instruments, we designed a response format that would allow us to examine various scoring options explored in previous self-efficacy research (Lee & Bobko, 1994). According to this format, the soldiers were presented with a stimulus (e.g., a leadership problem) and then were asked to respond considering their perceived or expected performance relative to that stimulus. First, soldiers were to indicate whether or not (*yes/no*) they believed that they could perform at a certain level. Second, the soldiers indicated how confident they were of their response on a nine point scale (1 = *Not Sure*; 9 = *Very Sure)*. An example item using this format is provided in Figure 6.2. They were to do this for five different levels of performance. Using this response format, the following scoring algorithms discussed by Lee and Bobko (1994) could be calculated for items on each of the instruments:

- Magnitude - an indication of the level at which the respondent believes he/she might have success. This score is calculated by,
  - coding the yes/no response for each level (yes=1, no=0), and
  - summing these coded yes/no responses.

- Strength - an indication of the respondent's confidence in their yes/no response. This score is calculated by,
  - summing the confidence ratings for each level, and
  - setting the score to missing if confidence ratings are missing for any level.

- Composite 1 - a variation of the Strength score, where the focus is on the confidence of the "yes" responses only. This score is calculated by,
  - multiplying the coded "yes/no" response by the confidence rating for each level, and
  - summing these scores across levels.

- Composite 2 - a variation of Composite 1, in which the standardized confidence ratings are summed. This score accounts for inherent differences in confidence at different levels, and focuses on the confidence in yes responses only. This score is calculated by
  - standardizing the confidence scale scores so that the mean confidence rating for the sample at each level equals 0 and the standard deviation equals 1,
  - multiplying the coded yes/no response by the standardized confidence rating for each level, and
  - summing these standardized scores.

Figure 6.2
Example item using response format meant to elicit a variety of self-efficacy scores (Lee & Bobko, 1994).

Considering the questions in the Army Leadership Questionnaire, do you think you got the right answer to...			How sure are you?								
	Yes	No	Not Sure								Very Sure
1. At least 50% of them?	_____	_____	1	2	3	4	5	6	7	8	9
2. At least 75% of them?	_____	_____	1	2	3	4	5	6	7	8	9
3. At least 85% of them?	_____	_____	1	2	3	4	5	6	7	8	9
4. At least 95% of them?	_____	_____	1	2	3	4	5	6	7	8	9
5. All of them?	_____	_____	1	2	3	4	5	6	7	8	9

Table 6.9 provides descriptive statistics and Hoyt reliabilities for the four scoring algorithms on the PNC. Table 6.10 presents the correlations among the algorithm scores. Note that while the reliability estimates for the SE strength and SE Composite 1 scores were higher, a significantly larger number of soldiers had a SE magnitude score than a SE strength, SE Composite 1, or SE Composite 2 score. The difference in the n was substantial enough to choose the Magnitude

score for further analyses, even though the correlation between Magnitude and Strength scores shows the scores were certainly not redundant.

Table 6.9

Descriptive Statistics and Reliabilities for Various Self-Efficacy Scores: Perceived Number of Correct Answers (PNC).

Self-Efficacy Score	n	M	SD	Hoyt r_{XX}
SE Magnitude	641	2.98	1.19	.67
SE Strength	586	7.04	1.50	.77
SE Composite 1	577	22.84	10.24	.77
SE Composite 2	577	0.33	2.60	n/a

Note: To calculate r_{XX} for Composite 2, more than one stimulus is required. PNC pertains to only one stimulus-- performance on the ALQ.

Table 6.10

Correlations Among Various Self-Efficacy Scores: Perceived Number of Correct Answers (PNC).

Self-Efficacy Score	1	2	3	4
Magnitude	--			
Strength	.26	--		
Composite 1	.88	.54	--	
Composite 2	.26	.73	.68	--

n ranges from 604 to 641

Level of Difficulty Associated With Leadership Problems (LDLP)

Recall that the LDLP required a respondent to consider his/her confidence that he/she could deal effectively with each of five leadership problems. The response format for each of these five problems was the same as that for the PNC. Thus the same scoring options were available.

The descriptive statistics and reliabilities for the four scoring algorithms are shown in Table 6.11. Correlations among the scores are provided in Table 6.12. Note that the reliability

coefficients associated with each of the scores are essentially the same; also note that there are many more respondents with a Magnitude score than there are with a Strength, Composite 1, or Composite 2 score. Given these results, we chose to retain the Magnitude score for further analyses.

Table 6.11

Descriptive Statistics and Reliabilities for Various Self-Efficacy Scores: Level of Difficulty Associated With Leadership Problems.

Self-Efficacy Score	n	M	SD	r_{xx}
Magnitude	638	3.80	1.17	.86
Strength	593	7.93	1.13	.90
Composite 1	603	31.41	11.06	.89
Composite 2	603	0.59	2.43	.89

Table 6.12

Correlations Among Various Self-Efficacy Scores: Level of Difficulty Associated With Leadership Problems.

Self-Efficacy Score	1	2	3	4
SE magnitude	--			
SE strength	.35	--		
SE Composite 1	.95	.58	--	
SE Composite 2	.35	.85	.62	--

n ranges from 619 to 641.

In conducting the validation study, we found that subjects had a difficult time understanding exactly how to complete the PNC and LDLP ratings. They had this difficulty despite the fact that we included an example in the instructions, we verbally explained how to complete the instruments, and we reminded them how to complete the instruments during testing. Our conclusion is that the response format required to produce the four cores discussed in previous research (Lee & Bobko, 1994) is inherently difficult for some subjects. Given these findings, we recommend simplifying the scoring format to provide the yes-no based magnitude scores only.

Correlations among Predictors

Table 6.13 reports correlations among each of the experimental predictor instruments and the non-experimental predictor instruments. The ALQ and LPI scores shows a moderate relationship (.32), but the relationships between scores on these two measures and the self-efficacy measures were all .17 or less. The three self-efficacy scores correlated .17, .18 and .27. In general, these measures were relatively distinct from one another.

Table 6.13

Correlations Between Experimental Predictors and Non-Experimental Predictors.

Predictor	1	2	3	4	5
Experimental Predictors					
1. ALQ	--				
2. LPI	32	--			
3. SE - ALQ	09	09	--		
4. PNC	10	07	18	--	
5. LDLP	17	17	17	27	--
ASVAB					
6. AFQT	27	18	04	08	11
7. Technical Factor	31	10	02	12	16
8. Quantitative Factor	27	16	-01	03	07
9. Verbal Factor	36	17	-01	11	15
10. Speed Factor	01	11	-00	-03	-01
Other Cognitive Markers					
11. Word Knowledge	35	19	11	10	19
12. Arithmetic Reasoning	26	18	07	06	15
13. Assembling Objects	22	15	01	07	11
ABLE					
14. Work Orientation	15	01	19	13	28
15. Adjustment	08	07	09	08	12
16. Dominance	14	02	21	17	27
17. Dependability	09	03	05	02	04
18. Locus of Control	09	01	08	12	12
19. Cooperation	-00	-02	05	-01	13
20. Physical Condition	-04	-04	03	07	05
21. Unlikely Virtues	-12	-09	-00	-04	03
22. Non-Random Response	24	25	12	01	08
Personnel File Form					
23. Recognition	-02	-03	06	10	12
24. Disciplinary Action	-02	10	04	-04	-00
25. Rec for Acc Promo	16	13	08	07	13

Note. Decimal points omitted; "--" = 1.0; n ranges from 430 to 670.

The ALQ showed a moderate, positive relationship with virtually all measures of cognitive abilities. It also showed a low positive relationship with the ABLE scores for Work Orientation and Dominance. This makes sense, in that the task required of those taking the ALQ includes aspects of problem solving (affected by g), attitudes toward work (affected by work orientation), and consideration of how one would interact in social situations (affected by dominance). We expected a similar pattern of correlations between the LPI and the various non-experimental predictors, but only a low positive relationship with the various measures of cognitive ability emerged. Both the ALQ and the LPI showed a strong positive relationship with the Non-Random Response score on the ABLE -- perhaps reflecting that participants who attended well to the test-taking activity performed better.

All three self-efficacy measures had generally low, positive relationships with cognitive test scores. They showed low to moderate relationships with both Work Orientation and Dominance scale scores from the ABLE. This makes sense, in that soldiers with a strong work orientation might develop greater self-efficacy, and those with greater self-efficacy may tend to dominate.

Criteria Analyses

As described in chapter 3, we developed four performance measures including the Structured Interview, the supervisory BARS (Behaviorally Anchored Rating Scales), the supervisory Situational Prediction Ratings, and the Personnel File Form (PFF). These analyses include only those NCOs who took all of those instruments. Some of these NCOs did not take the predictor measures or could not be matched with their predictor data because of incorrect SSNs.

Structured Interview

The analyses for the structured interview consisted of (a) descriptive statistics, (b) item correlations, and (c) internal consistency reliability. No interrater agreement analyses were performed because each NCO was rated by only one person. In the pilot study, the reliability of one rater was .77.

The descriptive statistics for the items are shown in Table 6.14. The internal consistency reliability of the interview, estimated using coefficient alpha, was .83 based on a sample size of 663. The item intercorrelations ranged from .42 to .56 as shown in Table 6.15.

Supervisory Ratings: BARS

All NCOs in the analyses had their performance rated by either a first-level supervisor, a second-level supervisor, or a peer. Attempts were made to use a first-level supervisor whenever possible, but many NCOs were rated by only a second-level supervisor or a peer. To assess interrater reliability, 53 NCOs were rated by two people. Once the interrater reliability analyses had been done,

Table 6.14
Descriptive Statistics for the Structured Interview.

Item	M	SD
1. Motivating Others	4.16	1.17
2. Demonstrating Effort	4.47	1.16
3. Planning and Providing for Training	4.19	1.07
4. Organizing, Coordinating, and Executing	4.11	1.09
5. Communicating Orally	4.60	1.13

Notes. N = 663. The rating scale for each item ranged from 1 to 7.

Table 6.15
Item Intercorrelations for the Structured Interview.

Item	1	2	3	4	5
1. Motivating Others	--				
2. Demonstrating Effort	55	--			
3. Planning and Providing for Training	52	49	--		
4. Organizing, Coordinating, and Executing	49	56	50	--	
5. Communicating Orally	46	42	43	50	--

Notes. N = 663. Decimal points omitted.

all subsequent analyses used only one rater for each NCO. When an NCO was rated by more than one person, a single rater was picked for the analyses according to the following preference order

1. First-level supervisor
2. Peer
3. Second-level supervisor
4. Length of time the rater has worked with or supervised the NCO.

For example, if one of the raters was a first-level supervisor and the other was a peer, the first-level supervisor was used; if both raters were at the same supervisory level relative to the NCO, then the one who had worked with or supervised the NCO longer was used.

The analyses for the supervisory BARS consisted of (a) descriptive statistics, (b) item correlations, (c) internal consistency reliability, and (d) interrater reliability.

The descriptive statistics for the items are shown in Table 6.16. The internal consistency reliability of the BARS, estimated using coefficient alpha, was .92 based on a sample size of 631. The item intercorrelations ranged from .28 to .64 as shown in Table 6.17.

The interrater reliability analyses for the BARS were based on the 53 NCOs who were rated by more than one person. If an NCO was rated by more than two people, then two of the ratings were randomly selected.

Table 6.16
Descriptive Statistics for the Supervisory BARS Ratings.

Item	M	SD
A. Demonstrating Technical Knowledge and Skill	5.16	1.27
B. Communicating Orally	4.87	1.35
C. Writing	4.51	1.23
D. Demonstrating Effort and Initiative	5.15	1.44
E. Following Regulations, Policies, and Procedures	5.38	1.24
F. Demonstrating Integrity and Discipline	5.54	1.24
G. Relating and Cooperating with Others	5.18	1.29
H. Motivating Others	4.70	1.46
I. Planning and Providing for Training	4.61	1.35
J. Directing, Monitoring, and Supervising Work	4.88	1.38
K. Organizing, Coordinating, and Executing	4.90	1.30
L. Demonstrating Responsiveness	4.95	1.36
M. Representing	5.14	1.26

Notes. N = 626. The rating scale for each item ranged from 1 to 7.
This procedure underestimates interrater reliability as it is commonly defined because reliability usually attributes to error only differences in the score patterns of the two raters whereas ICC(1,1) also attributes to error differences in the mean ratings of the two raters. Interrater reliability, as it is usually defined, could not be assessed because its computation would require that each person rate all of the NCOs. (In fact, most people rated only a few NCOs.)

Interrater reliability was estimated as the median correlation between the item score profiles for the two raters. There is one correlation for each NCO; the median correlation is across the NCOs. This estimates the amount of agreement among two raters on the dimension profiles.. The median and mean correlations were both .15.

A more conservative estimate of interrater reliability was computed using the intraclass correlation. The formula used was Shrout and Fleiss's (1979) ICC(1,1). The procedure for calculating ICC(1,1) is as follows (Shrout & Fleiss, 1979):

- Set up the data set so that there is a different case for each rater on each NCO.
- Compute a one-way ANOVA with NCO as the between-subjects factor and the rating as the dependent variable.

- True variance is that attributable to the NCO factor; error variance is that attributable to differences in ratings within an NCO.

Table 6.17

Item Intercorrelations for the Supervisory BARS Ratings.

Item	A	B	C	D	E	F	G	H	I	J	K	L	M
A. Technical Knowledge	--												
B. Oral Communication	49	--											
C. Writing	38	54	--										
D. Effort & Initiative	59	47	41	--									
E. Following Regs	43	43	39	56	--								
F. Integrity & Discipline	39	38	33	53	62	--							
G. Relating with Others	39	39	28	48	49	50	--						
H. Motivating Others	54	53	36	56	45	43	48	--					
I. Planning for Training	50	45	41	55	44	39	37	57	--				
J. Supervising Work	53	46	34	57	51	45	40	59	52	--			
K. Organizing	56	48	45	62	51	46	45	56	59	64	--		
L. Responsiveness	52	41	36	53	46	42	45	54	48	48	58	--	
M. Representing	36	41	34	53	51	57	51	48	41	47	46	44	--

Notes. N = 626. Decimals omitted. The item labels are abbreviated; see Table 6.16 for the full item labels.

Table 6.18 shows ICC(1,1) for each item in the supervisory BARS. The values ranged from -.06 to .40 with a median of .16. Although these values are low, they likely substantially underestimate interrater reliability as it is commonly defined. For the total BARS score, the interrater reliability was .20.

Although the interrater reliability estimates are somewhat lower than usually encountered (Campbell, Ford, Rumsey, Pulakos, Borman, Felker, DeVere, & Rieglehaupt , 1990), it must be kept in mind that NCO's are themselves supervisors and perform much of their job unobserved by their peers and superiors. This contributes substantially to the difficulty in obtaining agreement across raters.

Table 6.18

Interrater Agreement (ICC(1,1)) for the Supervisory BARS Ratings.

Item	ICC(1,1)	$p < .05$
A. Demonstrating Technical Knowledge and Skill	.17	
B. Communicating Orally	.00	
C. Writing	.34	.009
D. Demonstrating Effort and Initiative	.16	
E. Following Regulations, Policies, and Procedures	.11	
F. Demonstrating Integrity and Discipline	.40	.002
G. Relating and Cooperating with Others	.16	
H. Motivating Others	.17	
I. Planning and Providing for Training	.16	
J. Directing, Monitoring, and Supervising Work	-.03	
K. Organizing, Coordinating, and Executing	-.06	
L. Demonstrating Responsiveness	.04	
M. Representing	.20	
BARS TOTAL SCORE	.20	

Notes. N = 53. Shrout and Fleiss's (1979) ICC(1,1) was used to estimate interrater reliability. The column labeled $p < .05$ shows the significant p-values for the null hypothesis that the reliability is not greater than zero.

Supervisory Ratings: Situational Performance

The NCOs were rated using the situational performance questionnaire immediately after being rated using the BARS. The description of the data preparation for the BARS discussed above also applies to the situational ratings.

The analyses for the supervisory situational ratings consisted of (a) descriptive statistics, (b) item correlations, (c) internal consistency reliability, and (d) interrater reliability.

The descriptive statistics for the items are shown in Table 6.19. The internal consistency reliability of the supervisory situational ratings, estimated using coefficient alpha, was .92 based on a sample size of 629. The item intercorrelations ranged from .52 to .71 as shown in Table 6.20.

Table 6.19

Descriptive Statistics for the Supervisory Situational Ratings.

Item	M	SD
1. Handling Routine Requests or Complaints	4.91	1.25
2. Handling Conduct Problems	4.92	1.31
3. Handling a Shortage of Resources	4.80	1.39
4. Handling Neglect of Responsibilities	4.83	1.44
5. Handling Personal Problems	5.28	1.17
6. Handling Complaints about Command Actions	4.80	1.31
7. Handling Skill Deficits	4.97	1.35
8. Handling Subordinate Recognition or Promotion	5.17	1.29

Notes. N = 591. The rating scale for each item ranged from 1 to 7.

Table 6.20

Item Intercorrelations for the Supervisory Situational Ratings.

Item: Handling . . .	1	2	3	4	5	6	7	8
1. Routine Requests and Complaints	--							
2. Conduct Problems	71	--						
3. Shortage of Resources	63	57	--					
4. Neglect of Responsibilities	60	64	57	--				
5. Personal Problems	60	61	53	53	--			
6. Complaints about Command Actions	62	61	63	61	60	--		
7. Skill Deficits	58	54	65	55	52	58	--	
8. Subordinate Recognition or Promotion	63	59	61	57	60	63	61	--

Notes. N = 591. Decimals omitted.

Interrater reliability was estimated as the median correlation between the item score profiles for the two raters. There is one correlation for each NCO; the median correlation is across the NCOs. This estimates the reliability of one rater. The median correlation was .18 and the mean was .16.

As with the BARS, the intraclass correlation referred to by Shrout and Fleiss (1979) as ICC(1,1) was used to estimate the interrater reliability. Table 6.21 shows ICC(1,1) for each item in the supervisory situational ratings. The values ranged from -.27 to .19 with a median of .16. Although these values are low, they likely substantially underestimate interrater reliability as it is commonly defined. For the total score, the interrater reliability was .10. Thus, ICC(1,1) was lower for the supervisory Situational ratings than for the BARS. As with the BARS, the lack of opportunity to observe NCO performance likely affects the reliability estimates for these ratings.

Table 6.21

Interrater Agreement (ICC(1,1)) for the Supervisory Situational Ratings.

Item	ICC(1,1)	$p < .05$
1. Handling Routine Requests or Complaints	.19	
2. Handling Conduct Problems	.13	
3. Handling a Shortage of Resources	-.27	
4. Handling Neglect of Responsibilities	.24	.045
5. Handling Personal Problems	.03	
6. Handling Complaints about Command Actions	-.05	
7. Handling Skill Deficits	.14	
8. Handling Subordinate Recognition or Promotion	.02	
SITUATIONAL RATINGS TOTAL SCORE	.10	

Notes. N = 53. Shrout and Fleiss's (1979) ICC(1,1) was used to estimate interrater reliability. The column labeled $p < .05$ shows the significant p-values for the null hypothesis that the reliability is not greater than zero.

Personnel File Form (PFF)

The analyses for the Personnel File Form consisted of (a) descriptive statistics and (b) item correlations. Because this is a self-report instrument, interrater reliability did not apply.

The following scales were computed based on the PFF items:

- Awards: The number of awards, badges, or commendations received. Some items were given more weight than others when computing this variable. The weights were the same as those used in the Career Force study.
- Army Achievement Medals
- Army Commendation Medals
- Good Conduct Medals

- Letters: The number of letters/memoranda of appreciation, commendation, or achievement received while in grades E-5 or above.

- Certificates: The number of certificates of appreciation, commendation, or achievement received while in grades E-5 or above.

- Accelerated Promotions: The number of times the NCO received an accelerated promotion from an E-5 or higher grade.

- Physical Readiness Score: The possible score ranges from 0 to 300.

- Weapons Qualification: This was scored 1 for Marksman, 2 for Sharpshooter, and 3 for Expert.

- Articles 15: The number of Articles 15 received while in grades E-5 or above.

- Flag Actions: The number of Flag Actions received while in grades E-5 or above.

- Disciplinary Actions: The number of Articles15 or Flag Actions received while in grades E-5 or above.

- Awards, Letters, or Certificates: The sum of the number of memoranda/letters while in E-5 above, certificates while in E-5 or above, awards (as computed above), and medals received (see the Personnel File Form for more information).

The descriptive statistics for the items are shown in Table 6.22. The frequency distribution for each item was examined for skewness and kurtosis to determine if any transformations should be performed to make the data better approximate a normal distribution. Most of the variables did not exhibit severe skewness or kurtosis. The variables that deviated substantially from the normal distribution (Awards and the three variables related to disciplinary actions) could not be transformed because they consisted of essentially only two values. Therefore, no transformations were performed.

The magnitudes of the item intercorrelations ranged from .00 to .78 as shown in Table 6.23. Most of the correlations were quite low. The moderate and high correlations between Good Conduct Medals, Certificates, and Letters, however, show that these three scales are related.

Instrument Intercorrelations

The correlations between the total scores of the instruments were computed. Because the PFF does not have a total score, the correlations of each of its scales with the other measures was computed (see Table 6.23). Table 6.23 shows that the PFF variables are unrelated or have low, positive relationships to the other measures; all correlations are below .22.

Table 6.22

Descriptive Statistics for the PFF Scales.

Scale	n	M	SD	Skewness	Kurtosis
Awards	644	2.24	1.97	1.79	4.24
Army Achievement Medals	644	3.22	1.46	-0.35	-0.92
Army Commendation Medals	644	2.02	1.49	0.38	-0.83
Good Conduct Medals	644	3.15	1.46	-0.23	-1.01
Letters	644	2.95	2.39	0.36	-1.08
Certificates	644	3.03	2.34	0.35	-1.02
Accelerated Promotions	644	0.65	0.84	1.14	0.45
Physical Readiness Score	604	262.00	26.8	-0.60	-0.13
Weapons Qualification	620	2.52	0.68	-1.11	-0.06
Articles 15	644	0.10	0.33	3.52	12.70
Flag Actions	644	0.13	0.38	3.83	21.57
Disciplinary Actions	644	0.23	0.57	3.06	10.93
Letters, Awards, or Certificates	644	8.21	5.10	0.47	-0.51

Notes: Skewness values less than zero indicate that the skinny half of the distribution is to the left (i.e., most people have larger values of the variable). Kurtosis values less than zero indicate that the distribution is flatter than a normal distribution.

Table 6.24 shows that the two supervisory rating measures are highly related ($r = .83$). This is not surprising considering that the supervisors completed these two measures back-to-back. The structured interview, in contrast, correlates very little with the supervisory ratings scales. This low relationship may be due to the very different methods used in the supervisory ratings vs. the interview. In the interview, an SME rated the self-reports of the NCOs. The two instruments differed in another way, as well. The supervisory ratings involved the NCOs' typical performance whereas the SMEs rated the NCOs' best reported performance. That is, during interviews, interviewees tend to describe their best performances. Thus, the interview may have measured what the NCOs can do whereas the supervisors rated what the NCOs will do. That is, the supervisory ratings may include an important aspect of performance that the interview omits: motivation.

In addition, the interview is fakeable to some extent. Thus, to the extent that it was faked, the interview measured the NCOs' knowledge of what behaviors constitute good performance. Research on the faking of personality and attitude tests shows that although few people outright

lie, many shade the truth by always giving themselves the benefit of the doubt (Guion & Gottier, 1966).

Table 6.23
Scale Intercorrelations for the PFF.

Scale	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Awards	--												
2. Achievement Medals	17	--											
3. Commendation Medals	14	20	--										
4. Conduct Medals	12	20	49	--									
5. Letters	14	24	37	48	--								
6. Certificates	11	27	40	44	78	--							
7. Acceler. Promotions	07	09	22	12	27	28	--						
8. Physical Score	13	05	-06	10	10	12	01	--					
9. Weapons Qualification	04	09	-06	-08	04	02	05	14	--				
10. Articles 15	06	05	-01	12	12	08	01	-03	01	--			
11. Flag Actions	-02	02	-04	06	-00	01	01	-16	-09	28	--		
12. Disciplinary Actions	02	04	-03	11	07	05	01	-12	-06	77	83	--	
13. Letters/Awards/Certs.	50	30	40	47	88	86	28	15	04	11	-01	06	--
INTERVIEW TOTAL	05	16	10	05	21	15	19	07	08	01	-00	01	19
BARS TOTAL	11	14	16	09	19	15	18	17	06	-05	-04	-06	20
SITUATIONAL TOTAL	11	15	15	14	15	14	17	15	03	-00	02	01	18

Notes. Pairwise deletion: N varies from 601 to 644. Decimals omitted. Full names of scale labels are shown in the previous table. Scales 12 and 13 are totals from other scales: The correlations of these two scales with their constituent scales are underlined (and are therefore inflated because the constituent scales are part of the total score).

Table 6.24
Criteria Instrument Intercorrelations.

Instrument	Supervisory BARS	Supervisory Situational	Structured Interview
Supervisory BARS	--		
Supervisory Situational	83	--	
Structured Interview	16	09	--

Notes. Pairwise deletion: N = 611 to 631. Decimals omitted.

Development of a Single Criterion Variable

For the purposes of most of the validity analyses, a single criterion variable was developed. This approach was taken because it is most practical for selection and staffing purposes to have a single performance criterion, and it was not feasible within the constraints of project resources to conduct multiple sets of validation analyses.

Confirmatory factor analyses that included any PFF variables along with the other criteria variables would not run using LISREL. This is not surprising considering the low correlations of the PFF variables with the other instruments (see Table 6.23). Therefore, the PFF variables were excluded from the computation of the single criterion score. All the variables from the other three instruments (interview, supervisory BARS, supervisory Situational Ratings) were included in the computation of the single criterion score.

The following procedures were performed to compute the single criterion score for each NCO. First, each NCO's score for each of the three instruments was computed. Then these scores were transformed, within each instrument, to a distribution with a mean of 50 and a standard deviation of 10 (i.e., a T-score). Finally, these three instrument T-scores were summed for each NCO. This was the single criterion score. Thus, each of the three instruments was given equal weight. An NCO's criterion score was set to missing if more than 30% of the values in any of the three instruments were missing.

Validity Analyses

This section describes the analyses of the relationships between the predictor and criterion measures -- the validities. These analyses include:

- simple, uncorrected bivariate correlations between the market and new predictors and the criteria measures,
- multiple regressions corrected for range restriction and adjusted for shrinkage, used to estimate incremental validities of each new predictor, and
- multiple regressions, corrected and adjusted, used to estimate the absolute and incremental validity of the "optimum" or best set of new predictors.

Bivariate Correlations

Uncorrected, bivariate correlations between the single, overall criterion, the three major criterion instrument scores, and each of the predictors are provided in Table 6.25. Examination of the correlations between the predictors and the criterion instruments reveals no discernible differences in the pattern of results aside from the fact that correlations with the single criterion tend to be higher than those with the individual criterion instruments. This is likely due to the increased reliability and comprehensiveness in the criterion as it is the sum of the three instrument scores. Therefore, it appears unlikely that validation analyses conducted using the

three, separate criterion instrument scores would reveal anything different from analyses using the single, composite criterion.

Table 6.25
Correlations Between the Predictors and the Criteria.

Measure		1	2	3	4
Criteria					
1.	Criterion—Composite	--			
2.	Interview	56	--		
3.	Supervisor BARS Ratings	88	17	--	
4.	Supervisor Situational Ratings	84	10	83	--
Experimental Predictors					
5.	ALQ	20	13	15	15
6.	LPI	11	06	10	08
7.	SE - ALQ	09	05	08	06
8.	PNC	10	13	08	05
9.	LDLP	21	16	17	12
ASVAB					
10.	AFQT	11	12	07	03
11.	Technical Factor	12	12	08	05
12.	Quantitative Factor	14	09	12	08
13.	Verbal Factor	14	16	08	04
14.	Speed Factor	-02	04	-03	-05
Other Cognitive Marker Tests					
15.	Word Knowledge	12	15	09	04
16.	Arithmetic Reasoning	10	09	08	05
17.	Assembling Objects	06	06	06	01
ABLE					
18.	Work Orientation	36	29	26	21
19.	Adjustment	07	05	07	03
20.	Dominance	31	31	24	18
21.	Dependability	01	03	01	-01
22.	Locus of Control	16	11	12	14
23.	Cooperation	03	09	00	-04
24.	Physical Condition	06	11	04	00
25.	Unlikely Virtues	04	01	04	02
26.	Non-Random Response	08	01	09	06
Personnel File Form					
27.	Recognition	24	19	18	17
28.	Disciplinary Action	-02	01	-07	02
29.	Rec for Acc Promo	05	14	-03	-01

Note. Decimal points omitted; "--" = 1.0; n ranges from 380 to 652.

We also examined the bivariate correlations between the predictors and various PFF scores that correspond to performance after becoming an NCO, and thus might have served as

criteria: a) Count of Letters and Certificates of Appreciation, b) Count of Disciplinary Actions, c) Recommendations for Accelerated Promotions, d) Physical Readiness Test Score, and e) Weapons Score. Virtually all of these correlations were lower than the correlation between the given predictor and the overall criterion score. This makes sense given the more global nature of the criterion, the care with which the criterion score was developed, and the many non-performance oriented influences on the PFF scores.

Although these bivariate validities are not high, they are generally positive. The ALQ and LDLP show correlations of .20 and .21 with the overall criterion—higher than any of the ASVAB or other cognitive marker test correlations. The ABLE shows relatively impressive relationships for two of its scales, Work Orientation and Dominance. Finally, the Recognition score from the PFF shows a correlation with the criterion of .24, higher than all but the two ABLE scale correlations.

Regression Analyses to Assess Incremental Validity of Experimental Predictors

To assess the incremental validity of each experimental predictor over the non-experimental predictors, we conducted several sets of regression analyses. Each set of regression analyses followed the pattern shown in Table 6.26.

Table 6.26

Pattern of Regressions Used to Evaluate Incremental Validity of Experimental Predictors.

Regression	Predictors in Equation
1	Experimental Predictor
2	Cognitive Ability
3	Cognitive Ability, Experimental Predictor
4	Cognitive Ability, ABLE
5	Cognitive Ability, ABLE, Experimental Predictor
6	Cognitive Ability, ABLE, PFF
7	Cognitive Ability, ABLE, PFF, Experimental Predictor

We designed this series of regressions to show the incremental validity of the experimental predictors in various likely prediction scenarios. The first regression is meant to show the validity of the experimental measure when used alone in predicting performance. The second, fourth, and sixth regressions include only non-experimental predictors, and thus provide baseline validities for the third, fifth, and seventh regressions, each of which includes the experimental predictor. Each successive regression includes additional non-experimental predictors in order of probability of use. The second regression shows the validity of the measures of cognitive ability in this sample. This is important as it is likely that some form of cognitive ability score or some surrogate thereof might be used to predict NCO performance. The fourth regression includes another potentially useful predictor of NCO performance-- the

ABLE. This instrument is used as a predictor along with cognitive ability, as it is unlikely that it would be used any other way operationally. The sixth regression includes additional potentially useful predictors of NCO performance-- personnel file actions. These measures are less likely to be formally adopted as predictors of performance than are either cognitive ability scores or ABLE scores, so all of these predictors are included together in a regression equation. Thus, the third, fifth, and seventh regressions allow us to examine the incremental validity of the experimental predictor over a growing set of non-experimental predictors.

When examining the incremental validity of each experimental predictor, we conducted the first, and most important, set of regressions using the AFQT as the cognitive ability marker. We chose the AFQT because it is used widely and the score was available on nearly 90 percent of the sample. In addition, as the AFQT scores were normed to the 1980 youth population (DoD, 1982), we could use the AFQT to correct the regression covariance matrix for range restriction . Most of our discussion and conclusions are based on this first set of regressions. We conducted additional sets of regressions to rule out various alternative explanations for the findings.

In conducting the regression analyses, we engaged in the following steps:

- Identify all variables to be used in the set of regressions at hand.
- Retain only those observations that are nonmissing on all variables identified.
- Obtain the population covariance matrix for the variable(s) that will be used to correct for range restriction.
- Calculate the sample covariance matrix.
- Correct the sample covariance matrix for range restriction using the population covariance matrix.
- Conduct the set of regression analyses using the corrected covariance matrix.
- Adjust the Multiple R for shrinkage using Rozeboom's (1978) formula 8.

We corrected for range restriction using either the AFQT or the nine ASVAB test scores. The variance of the AFQT in the 1980 youth population was 785.68 (DoD, 1982). The covariance matrix of the ASVAB test scores in the 1980 youth population was also obtained from DoD (1982).

ALQ

The results for the first set of regressions are listed in Table 6.27. Notice that the ALQ increments the multiple regression by .02 above and beyond all the other predictors, and has a substantial amount of incremental variance beyond that predicted by a combination of the AFQT and Assembling Objects Test scores (.16), a moderate correlation with the criterion. Notice also that the absolute level of prediction obtained by the full-equation (.39) shows considerable utility.

Table 6.27
Multiple Regression Results Using AFQT as the Cognitive Ability Marker, Experimental Predictor: ALQ.

Basic Prediction Equation[a]	Basic Equation			Basic Equation+ALQ			
	R	R_C	R_{CA}	R	R_C	R_{CA}	ΔR
Y=ALQ	.22	.22	.21				
Y=AFQT	.08	.10	.07	.22	.23	.21	.14
Y=AFQT+AO	.09	.10	.03	.22	.23	.19	.16
Y=AFQT+AO+ABLE	.41	.41	.37	.43	.44	.39	.02
Y=AFQT+AO+ABLE+PFF	.43	.43	.37	.45	.45	.39	.02

Note: n = 421. [a]Each regressor was allowed to take on an optimal b weight. R_C = R corrected for range restriction in AFQT scores; R_{CA} = R_C Adjusted for shrinkage; ABLE predictors include 7 scales (Work Orientation, Adjustment, Dominance, Dependability, Locus of Control, Cooperation, and Physical Condition). PFF predictors include 3 scores (Recognition, Disciplinary Actions, and Recommendations for Accelerated Promotions).

Table 6.28 presents the beta weights for the regressions. Notice that when all predictors are in the equation, the ALQ retains a substantial beta weight. Other predictors that contribute substantially to the prediction of the criterion include the Work Orientation, Dominance, and Dependability scale scores from the ABLE, and the number of Recommendations for Accelerated Promotions. The AFQT and Assembling Objects Test virtually fail to account for any unique variance in the criterion. These findings are not particularly surprising given that the criterion focuses on leadership performance almost to the exclusion of the technical aspects of MOS performance, where the technical aspects of MOS performance are likely to be those most related to cognitive ability (McHenry et al, 1990).

One somewhat puzzling finding here are the negative weights assigned to the Adjustment, Cooperativeness, and Dependability scale scores from the ABLE. These are counter-intuitive, and should be interpreted with considerable caution. It is highly likely that restricting these beta weights to zero or positive would result in very little or no loss in predictive power.

The results of this first set of regressions lend support to the notion that the ALQ predicts variance in NCO performance beyond that predicted by other predictors such as the AFQT or the ABLE. However, there are various alternative explanations for these results. Below, we address some of these through additional sets of regression analyses.

One potential criticism of the results is that some of the variance in the criterion predicted only by the ALQ might well be predicted by a more detailed set of ASVAB scores currently available to Army decision-makers. Therefore, our evaluation of the incremental validity of the ALQ would be incomplete unless we evaluated it with the ASVAB factors as a cognitive ability marker. We performed two more sets of regressions addressing the issue. We conducted the second set of regressions using the ASVAB factors as the cognitive marker, and using the nine ASVAB scores to perform a multivariate range restriction correction to the sample covariance matrix. There are 10 ASVAB subtests. Two are commonly combined to form one score. As

Table 6.28

Beta Weights From Multiple Regression Using AFQT as the Cognitive Ability Marker, Experimental Predictor: ALQ.

Predictors in Equation	R_{CA}	ΔR	β ALQ	AFQT	AO	ADJ	COO	DEP	DOM	LOC	PHY	WOR	REC	DA	PRO
ALQ	21		22												
AFQT	07			10											
AFQT, ALQ	20	14	21	03											
AFQT, AO	03			07	04										
AFQT, AO, ALQ	19	16	21	02	03										
AFQT, AO, ABLE	37			08	02	-08	-08	-09	20	07	-03	26			
AFQT, AO, ABLE, ALQ	39	02	15	05	01	-08	-08	-10	19	06	-03	25			
AFQT, AO, ABLE, PFF	37			05	01	-06	-08	-10	17	07	-03	25	06	-06	11
AFQT, AO, ABLE, PFF, ALQ	39	02	14	02	00	-06	-08	-10	16	06	-03	24	06	-06	10

Note: n = 421. Decimals omitted. R_{CA} = Multiple R, corrected for range restriction, adjusted for shrinkage; ΔR = change in R_{CA} with and without the ALQ in equation; ALQ = Army Leadership Questionnaire; AFQT = Armed Forces Qualification Test; AO = Assembling Objects. ABLE scale names abbreviated as follows: ADJ = Adjustment; COO = Cooperation; DEP = Dependability; DOM = Dominance; LOC = Locus of Control; PHY = Physical Condition; WOR = Work Orientation. PFF measures abbreviated as follows: REC = Recognition; DA = Disciplinary Action; PRO = Recommendation for Accelerated Promotions.

described earlier ASVAB factor scores were available for a subset of the full sample. Thus the sample size for the second analysis was reduced. Then we conducted a third set of regression analyses in a manner similar to that used for the first set of regressions (i.e., using AFQT as the cognitive ability marker and correcting for range restriction using the AFQT) with the same observations used in the second set of regressions. The results of these regressions are provided in Table 6.29

Table 6.29

Comparison of ALQ Incremental Validities When AFQT Scores or ASVAB Factor Scores are Used as Cognitive Ability Markers.

Basic Prediction Equation[a]	Basic Equation			Basic Equation+ALQ			
	R	R_C	R_{CA}	R	R_C	R_{CA}	ΔR
Cognitive Ability Marker: ASVAB Four Factors[b]							
Y=ALQ	.21	.24	.23				
Y=ASVAB4F	.14	.19	.10	.23	.26	.20	.09
Y=ASVAB4F+AO	.15	.19	.09	.23	.26	.19	.10
Y=ASVAB4F+AO+ABLE	.43	.44	.37	.45	.46	.39	.02
Y=ASVAB4F+AO+ABLE+PFF	.45	.46	.37	.47	.48	.39	.02
Cognitive Ability Marker: AFQT[c]							
Y=ALQ	.21	.23	.22				
Y=AFQT	.13	.17	.15	.22	.25	.22	.07
Y=AFQT+AO	.15	.18	.15	.23	.25	.22	.07
Y=AFQT+AO+ABLE	.43	.44	.38	.45	.46	.40	.02
Y-AFQT+AO+AB:E+PFF	.45	.46	.40	.46	.47	.40	.01

Note: n = 330. [a]Each regressor was allowed to take on an optimal b weight. R_C = R corrected for range restriction; [b]Corrections for range restriction were conducted using the nine ASVAB test scores. [c]Corrections for range restriction were conducted using the AFQT. R_{CA} = R_C Adjusted for shrinkage; ASVAB4F = ASVAB four factor scores; ABLE predictors include 7 scales (Work Orientation, Adjustment, Dominance, Dependability, Locus of Control, Cooperation, and Physical Condition). PFF predictors include 3 scores (Recognition, Disciplinary Actions, and Recommendations for Accelerated Promotions).

Note that the incremental validity of the ALQ is indeed lower in the second set of regressions using the four ASVAB factor scores, than it is in the first set of regressions, using the AFQT. However, also note that the results from the third set of regressions, using the AFQT scores, are virtually identical to those from the second set. Indeed, the incremental validates were greater over the ASVAB factor scores than over the AFQT. Thus, the lower incremental validities found in the second set of regressions are likely due to differing sample characteristics, not because the four factor scores were used.

Another criticism that may be leveled at the findings from the first set of regressions is that they may be due to the fact that the AFQT scores are old. Indeed, some of the scores are over 20 years old. Changes can occur over time that may affect the cognitive ability of study participants, thus leading to decay in the relevance of the AFQT scores, and a possible

overestimation of the incremental validities. To address this issue, we ran a fourth set of regressions using the Word Knowledge Test, the Arithmetic Reasoning Test, and the Assembling Objects Test as cognitive ability markers. We corrected the covariance matrix using the AFQT score. The results of this set of regressions are provided in Table 6.30. Note that the results are again virtually identical to those obtained with the first set of regressions. The incremental validity of the ALQ over the cognitive ability markers is again substantial, and there is a small incremental validity for the ALQ when the ABLE scores are also used to predict variance in performance. Thus, it appears that the demonstrated incremental validity of the ALQ is not due to a substantial amount of decay in the AFQT scores. These findings are in accord with those from Project A and Career Force, which showed little decrement in the validity of ASVAB scores over time (Rumsey, Peterson, Oppler, & Campbell, 1996).

Table 6.30
Evaluation of ALQ Incremental Validity With Cognitive Ability Measures Obtained Recently.

Basic Prediction Equation[a]	Basic Equation			Basic Equation+ALQ			
	R	R_C	R_{CA}	R	R_C	R_{CA}	ΔR
Y=ALQ	.22	.22	.21				
Y=WK+AR+AO	.09	.11	.00	.22	.23	.18	.18
Y=WK+AR+AO+ABLE	.41	.41	.35	.43	.43	.38	.03
Y=WK+AR+AO+ABLE+PFF	.43	.43	.37	.45	.45	.39	.02

Notes. n = 420. [a]Each regressor was allowed to take on an optimal b weight. R_C = R corrected for range restriction the AFQT; R_{CA} = R_C Adjusted for shrinkage; ABLE predictors include 7 scales (Work Orientation, Adjustment, Dominance, Dependability, Locus of Control, Cooperation, and Physical Condition). PFF predictors include 3 scores (Recognition, Disciplinary Actions, and Recommendations for Accelerated Promotions).

LPI

The results of the first set of regressions meant to evaluate the incremental validity of the LPI are listed in Table 6.31. Here again, this set of regressions is perhaps the most important, as the sample size is larger and more representative of the NCO population. Notice that the LPI predicts essentially no variance in the criterion-- even when it is the only variable in the prediction equation. The incremental R associated with the LPI for each of the equations is less than 0, as the LPI accounts for no variance, and the addition of the LPI variable to the prediction function contributes to shrinkage in the R.

Table 6.32 presents the beta weights for the regressions. Notice that even when the LPI is used alone in the prediction equation, the beta weight associated with LPI is near 0. The findings regarding the predictive contributions of the non-experimental markers are similar to those found with the ALQ regressions. The predictors that contribute substantially to the prediction of the criterion include the Work Orientation and Dominance scale scores from the ABLE, and the number of Recommendations for Accelerated Promotions. The AFQT and Assembling Objects Test appear to add little if anything to the prediction of the criterion.

Table 6.31
Multiple Regression Results Using AFQT as the Cognitive Ability Marker, Experimental Predictor: LPI.

Basic Prediction Equation[a]	Basic Equation			Basic Equation+LPI			
	R	R_C	R_{CA}	R	R_C	R_{CA}	ΔR
Y=LPI	.06	.06	.00				
Y=AFQT	.08	.10	.08	.09	.11	.05	-.02
Y=AFQT+AO	.09	.11	.05	.10	.12	.00	-.05
Y=AFQT+AO+ABLE	.41	.41	.37	.41	.41	.36	-.00
Y=AFQT+AO+ABLE+PFF	.43	.43	.37	.43	.43	.37	.00

Note: n = 428. [a]Each regressor was allowed to take on an optimal b weight. R_C = R corrected for range restriction in AFQT scores; R_{CA} = R_C Adjusted for shrinkage; ABLE predictors include 7 scales (Work Orientation, Adjustment, Dominance, Dependability, Locus of Control, Cooperation, and Physical Condition). PFF predictors include 3 scores (Recognition, Disciplinary Actions, and Recommendations for Accelerated Promotions).

To further assess the issue of the incremental validity of the LPI, we computed additional sets of regressions using different cognitive ability markers, as we had for the ALQ. For the second set of regressions, we used ASVAB factor scores as the cognitive ability marker, and the nine ASVAB test scores to correct for range restriction. For the third set of regressions, we used Word Knowledge, Arithmetic Reasoning, and Assembling Objects as cognitive ability markers, and the AFQT to correct for range restriction. The results of both sets of regressions are provided in Table 6.33. As in the first set of regressions, the LPI fails to predict any variance in the criterion in these regressions. None of these findings provide support for the notion that the LPI could predict NCO leadership performance.

Self-Efficacy

SE scores were retained as predictors: SE-ALQ, SE-PNC, SE-LDLP. We conducted regressions first including all 3 measures together. The results of the first set of regressions meant to evaluate the incremental validity of the self-efficacy measures are listed in Table 6.34. The self-efficacy measures increment prediction of the criterion by .01 above and beyond the combination of the AEQT, AO and ABLE, but do not increment prediction when the PFF is added. They show a considerable amount of incremental prediction beyond that obtained by a combination of the AFQT and Assembling Objects Test scores; and when used as the only predictors, the multiple R is one point higher than that found for the ALQ (.22).

Table 6.35 presents the beta weights for the regressions. Notice that the LDLP receives the highest beta weight of the three measures, and retains that relative standing when all predictors are in the equation. The other self-efficacy measures fail to contribute much to prediction even when only the self-efficacy measures are used in the prediction equation. As for the other regressions, the predictors that contribute substantially to the prediction of the criterion

Table 6.32

Beta Weights From Multiple Regressions Using AFQT as the Cognitive Ability Marker, Experimental Predictor: LPI.

			BETA WEIGHTS												
Predictors in Equation	R_{CA}	ΔR	LPI	AFQT	AO	ADJ	COO	DEP	DOM	LOC	PHY	WOR	REC	DA	PRO
LPI	00		06												
AFQT	08			10											
AFQT, LPI	05	-02	04	09											
AFQT, AO	05			08	04										
AFQT, AO, LPI	00	-05	04	07	04										
AFQT, AO, ABLE	36			09	02	-09	-09	-08	21	06	-03	27			
AFQT, AO, ABLE, LPI	36	00	05	08	02	-10	-09	-08	21	06	-03	26			
AFQT, AO, ABLE, PFF	37			06	01	-07	-09	-09	17	06	-03	25	06	-06	12
AFQT, AO, ABLE, PFF, LPI	37	00	05	05	01	-08	-09	-09	17	06	-03	25	07	-07	11

Notes. n = 421. Decimals omitted. R_{CA} = Multiple R, corrected for range restriction, adjusted for shrinkage; ΔR = change in R_{CA} with and without the ALQ in equation; ALQ = Army Leadership Questionnaire; AFQT = Armed Forces Qualification Test; AO = Assembling Objects. ABLE scale names abbreviated as follows: ADJ = Adjustment; COO = Cooperation; DEP = Dependability; DOM = Dominance; LOC = Locus of Control; PHY = Physical Condition; WOR = Work Orientation. PFF measures abbreviated as follows: REC = Recognition; DA = Disciplinary Action; PRO = Recommendations for Accelerated Promotions.

Table 6.33

Evaluation of LPI Incremental Validity With Different Cognitive Ability Markers.

Basic Prediction Equation[a]	Basic Equation			Basic Equation+LPI			
	R	R_C	R_{CA}	R	R_C	R_{CA}	ΔR
Cognitive Ability Marker: ASVAB Four Factors (n = 336)[b]							
Y=LPI	.05	.09	.04				
Y=ASVAB4F	.14	.19	.12	.14	.20	.09	-.02
Y=ASVAB4F+AO	.15	.20	.10	.15	.20	.08	-.03
Y=ASVAB4F+AO+ABLE	.43	.44	.37	.43	.44	.36	-.01
Y=ASVAB4F+AO+ABLE+PFF	.45	.46	.37	.45	.46	.37	-.01
Cognitive Ability Marker: Word Knowledge, Arithmetic Reasoning, Assembling Objects (n = 427)[c]							
Y=LPI	.06	.06	.00				
Y=WK+AR+AO	.09	.11	.00	.10	.12	.00	.00
Y=WK+AR+AO+ABLE	.40	.41	.35	.41	.41	.35	-.00
Y=WK+AR+AO+ABLE+PFF	.43	.43	.37	.43	.43	.36	-.00

Note: [a]Each regressor was allowed to take on an optimal b weight. R_C = R corrected for range restriction. [b]Corrections for range restriction were conducted using the nine ASVAB test scores. [c]Corrections for range restriction were conducted using the AFQT. R_{CA} = R_C Adjusted for shrinkage; ABLE predictors include 7 scales (Work Orientation, Adjustment, Dominance, Dependability, Locus of Control, Cooperation, and Physical Condition). PFF predictors include 3 scores (Recognition, Disciplinary Actions, and Recommendations for Accelerated Promotions).

Table 6.34

Multiple Regression Results Using AFQT as the Cognitive Ability Marker, Experimental Predictors: Self-Efficacy Scores.

Basic Prediction Equation[a]	Basic Equation			Basic Equation+SE			
	R	R_C	R_{CA}	R	R_C	R_{CA}	ΔR
Y=SE	.25	.26	.22				
Y=AFQT	.11	.14	.12	.26	.28	.24	.11
Y=AFQT+AO	.12	.14	.10	.26	.28	.23	.13
Y=AFQT+AO+ABLE	.43	.43	.38	.45	.45	.39	.01
Y=AFQT+AO+ABLE+PFF	.45	.45	.39	.47	.47	.40	.00

Notes. n = 374. [a]Each regressor was allowed to take on an optimal b weight. R_C = R corrected for range restriction in AFQT scores; R_{CA} = R_C Adjusted for shrinkage; SE scores include SE-ALQ, PNC, and LDLP; ABLE predictors include 7 scales (Work Orientation, Adjustment, Dominance, Dependability, Locus of Control, Cooperation, and Physical Condition). PFF predictors include 3 scores (Recognition, Disciplinary Actions, and Recommendations for Accelerated Promotions).

Table 6.35

Beta Weights From Multiple Regressions Using AFQT as the Cognitive Ability Marker, Experimental Predictor: SE.

Predictors in Equation	BETA WEIGHTS																	
	R_{CA}	ΔR	SE-ALQ	PNC	LDLP	AFQT	AO	ADJ	COO	DEP	DOM	LOC	PHY	WOR	REC	DA	PRO
SE	.22		-.01	.04	.24												
AFQT	.12					.14											
AFQT, SE	.24	.11	-.01	.03	.23	.10											
AFQT, AO	.10					.13	.02										
AFQT, AO, SE	.23	.13	-.01	.03	.23	.10	.00										
AFQT, AO, ABLE	.38					.14	-.01	-.10	-.08	-.10	.19	.05	-.01	.29			
AFQT, AO, ABLE, SE	.39	.01	-.05	-.03	.15	.13	-.01	-.11	-.08	-.09	.19	.05	-.01	.26			
AFQT, AO, ABLE, PFF	.39					.10	-.01	-.08	-.08	-.10	.16	.05	-.01	.27	.05	-.06	.13
AFQT, AO, ABLE, PFF, SE	.40	.00	-.05	-.03	.13	.09	-.02	-.09	-.08	-.10	.16	.05	-.01	.25	.04	-.06	.12

Notes. n = 374. Decimals omitted. SE scores include SE-ALQ, PNC, and LDLP; R_{CA} = Multiple R, corrected for range restriction, adjusted for shrinkage; ΔR = change in R_{CA} with and without the ALQ in equation; ALQ = Army Leadership Questionnaire; AFQT = Armed Forces Qualification Test; AO = Assembling Objects. ABLE scale names abbreviated as follows: ADJ = Adjustment; COO = Cooperation; DEP = Dependability; DOM = Dominance; LOC = Locus of Control; PHY = Physical Condition; WOR = Work Orientation. PFF measures abbreviated as follows: REC = Recognition; DA = Disciplinary Action; PRO = Recommendation for Accelerated Promotions

include the Work Orientation and Dominance scale scores from the ABLE, and the number of Recommendations for Accelerated Promotions. In contrast to the earlier reported regressions pertaining to the ALQ and the LPI, the AFQT contributes at least marginally to the prediction of the criterion beyond that of the other measures. This can likely be attributed to the fact that cognitive ability is likely to be relevant to performance on the ALQ and the LPI, but not to scores on the self-efficacy instruments. Indeed, a look at the correlations in Table 6.13 confirms this. The ALQ and the LPI correlate somewhat with the various cognitive markers, while the Self-efficacy scores generally do not.

The results of this first set of regressions lend support to the notion that the LDLP predicts variance in NCO performance beyond that predicted by other predictors such as the AFQT, and, to a lesser extent, the ABLE. As for the ALQ and LPI, there are similar alternative explanations, which we examined in the same way we did for the ALQ and LPI.

Table 6.36 provides the results of regressions intended to address the alternative explanations of a more detailed set of ASVAB scores.

Table 6.36

Comparison of Self-Efficacy Score Incremental Validities When AFQT Scores or ASVAB Factor Scores are Used as Cognitive Ability Markers.

Basic Prediction Equation[a]	Basic Equation			Basic Equation+SE			
	R	R_C	R_{CA}	R	R_C	R_{CA}	ΔR
Cognitive Ability Marker: ASVAB Four Factors [b]							
Y=SE	.26	.27	.23				
Y=ASVAB4F	.17	.27	.21	.28	.35	.28	.07
Y=ASVAB4F+AO	.17	.27	.20	.29	.35	.27	.07
Y=ASVAB4F+AO+ABLE	.44	.48	.41	.46	.50	.41	.00
Y=ASVAB4F+AO+ABLE+PFF	.46	.50	.41	.48	.51	.41	.00
Cognitive Ability Marker: AFQT [c]							
Y=SE	.26	.27	.23				
Y=AFQT	.17	.22	.20	.29	.32	.28	.08
Y=AFQT+AO	.17	.22	.19	.29	.32	.27	.08
Y=AFQT+AO+ABLE	.44	.46	.40	.46	.48	.41	.01
Y=AFQT+AO+ABLE+PFF	.46	.48	.40	.48	.49	.40	.00

Notes. n = 293. [a]Each regressor was allowed to take on an optimal b weight. R_C = R corrected for range restriction; [b]Corrections for range restriction were conducted using the nine ASVAB test scores. [c]Corrections for range restriction were conducted using the AFQT. R_{CA} = R_C Adjusted for shrinkage; ASVAB4F = ASVAB four factor scores; ABLE predictors include 7 scales (Work Orientation, Adjustment, Dominance, Dependability, Locus of Control, Cooperation, and Physical Condition). PFF predictors include 3 scores (Recognition, Disciplinary Actions, and Recommendations for Accelerated Promotions).

The incremental validity of the SE measures in Table 6.36 is indeed lower than it is in Table 6.34. But note that the results of the AFQT regressions in Table 6.36 are virtually identical to those for the four-factor score results. Thus, as with the earlier reported ALQ results, the lower incremental validities found in Table 6.36 are likely due to the different characteristics of this reduced sample rather than the predictive power of the more detailed cognitive ability battery. Note that $\underline{R}_{ca} = .28$ for both the ASVAB and AFQT in Table 6.36 (reduced sample) versus .24 for AFQT in Table 6.34 (larger sample). As the results in this reduced sample are the same regardless of the cognitive ability marker used, it is unlikely that incremental validities found in the larger sample would be greatly reduced if the ASVAB factors were available for all observations and used as the cognitive ability marker.

As already noted, the age of the AFQT scores may be raised as a criticism of the findings from the first set of regressions. To address this issue, we ran a set of regressions using the Word Knowledge Test, the Arithmetic Reasoning Test, and the Assembling Objects Test as cognitive ability markers. We corrected the covariance matrix using the AFQT score. The results of this set of regressions are provided in Table 6.37. Note that the results are again virtually identical to those obtained on Table 6.34 (compare row 1 in each table, and rows 3, 4 and 5 in Table 6.34 to rows 2, 3 and 4 in Table 6.37). The incremental validity of the SE predictors over the cognitive ability markers is again substantial, and there is a small incremental validity for the SE predictors when the ABLE scores are also used to predict variance in performance. Thus, it appears that the demonstrated incremental validity of the SE predictors is not due to a substantial amount of decay in the AFQT scores.

Table 6.37

Evaluation of SE Measures: Incremental Validity Over Cognitive Ability Measures Obtained Recently

Basic Prediction Equation[a]	Basic Equation			Basic Equation+SE			
	$\underline{R}$	$\underline{R}_C$	$\underline{R}_{CA}$	$\underline{R}$	$\underline{R}_C$	$\underline{R}_{CA}$	Δ
Y=SE	.25	.26	.22				
Y=WK+AR+AO	.12	.15	.08	.26	.27	.21	.13
Y=WK+AR+AO+ABLE	.42	.43	.37	.44	.45	.38	.01
Y=WK+AR+AO+ABLE+PFF	.45	.45	.38	.46	.47	.39	.01

Note: $\underline{n} = 374$. [a]Each regressor was allowed to take on an optimal $\underline{b}$ weight. $\underline{R}_C = \underline{R}$ corrected for range restriction the AFQT; $\underline{R}_{CA} = \underline{R}_C$ Adjusted for shrinkage; ABLE predictors include 7 scales (Work Orientation, Adjustment, Dominance, Dependability, Locus of Control, Cooperation, and Physical Condition). PFF predictors include 3 scores (Recognition, Disciplinary Actions, and Recommendations for Accelerated Promotions).

Examination of each of the standardized regression weights associated with each set of regressions shows that the LDLP consistently contributes to the prediction of the criterion, even when other predictors such as the AFQT and the ABLE are included as predictors in the

equation. The other self-efficacy measures fail to predict variance in the criterion even when the self-efficacy measures are the only predictors in the regression equation. Table 6.38 presents the regression results examining the incremental validity of the LDLP alone. Note that the multiple Rs and incremental validities in these regressions are virtually identical to those presented earlier when the PNC and SE-ALQ scores were also included in the regressions.

Table 6.38

Multiple Regression Results Using AFQT as the Cognitive Ability Marker, Experimental Predictor: LDLP.

Basic Prediction Equation[a]	Basic Equation			Basic Equation+LDLP			
	R	R_C	R_{CA}	R	R_C	R_{CA}	ΔR
Y=LDLP	.23	.23	.22				
Y=AFQT	.09	.11	.09	.24	.25	.23	.14
Y=AFQT+AO	.09	.11	.05	.24	.25	.22	.16
Y=AFQT+AO+ABLE	.41	.42	.37	.43	.43	.38	.01
Y=AFQT+AO+ABLE+PFF	.43	.44	.38	.44	.45	.39	.01

Note: n = 414. [a]Each regressor was allowed to take on an optimal b weight. R_C = R corrected for range restriction in AFQT scores; R_{CA} = R_C Adjusted for shrinkage; ABLE predictors include 7 scales (Work Orientation, Adjustment, Dominance, Dependability, Locus of Control, Cooperation, and Physical Condition). PFF predictors include 3 scores (Recognition, Disciplinary Actions, and Recommendations for Accelerated Promotions).

These findings lend support to the notion that the LDLP could be used to predict NCO performance above and beyond that predicted by cognitive ability measures. These findings also lend support to the more traditional methods of measuring self-efficacy, wherein a respondent is to rate his/her confidence in performing a variety of tasks (Bandura, 1982).

These findings fail to support the use of SE-ALQ or PNC in predicting NCO performance. These instruments were designed to offset potential fakeability—a limitation of the LDLP. However, the methods used proved not valid in predicting performance. It may be that a score that includes both SE-ALQ and actual ALQ performance might be valid.

Regression Analyses to Assess Optimum Equation

To this point, all of the regression analyses that we have reported were done to evaluate the incremental validity of individual experimental predictors. Having identified two experimental predictors with substantial incremental validity, we now turn to evaluating the total incremental validity that can be attributed to these predictors. To address this issue, we conducted a set of regressions following the pattern shown in Table 6.26. The experimental

predictors in these regressions were the ALQ and the LDLP. In addition, only those non-experimental predictors that appeared to contribute to prediction of the criterion in earlier regressions were included. Results of these regressions are provided in Table 6.39. In this sense, we called these "optimum" equations since they included only those predictors considered likely to be most valid, based on earlier findings.

Table 6.39
Multiple Regression Results Using AFQT as the Cognitive Ability Marker, Experimental Predictors: ALQ+LDLP.

Basic Prediction Equation[a]	Basic Equation			Basic Equation+ALQ+LDLP			
	R	R_C	R_{CA}	R	R_C	R_{CA}	ΔR
Y=ALQ+LDLP	.29	.30	.28				
Y=AFQT	.08	.10	.07	.29	.30	.28	.21
Y=AFQT+DEP+DOM+WOR	.40	.40	.38	.43	.44	.41	.03
Y=AFQT+DEP+DOM+WOR+ACCP	.41	.41	.39	.44	.45	.41	.02

Note: n = 409. [a]Each regressor was allowed to take on an optimal b weight. R_C = R corrected for range restriction in AFQT scores; R_{CA} = R_C Adjusted for shrinkage.

Together the ALQ and LDLP increment prediction of the criterion above and beyond all the other predictors by about .02. They increment prediction over the AFQT by a substantial amount (.21); and when used as the only predictors, the multiple R is .28. Also note that the ALQ and LDLP compliment one another in the prediction of the criterion—the multiple Rs and incremental validities are higher in these regression results than they are in either the ALQ or LDLP results presented earlier (about .03 to .07 or higher). This is the case even after adjusting for shrinkage.

Table 6.40 presents the beta weights for the regressions. Notice that when all predictors are in the equation, the ALQ retains a substantial beta weight (.15), about 25% less than its weight when entered with just LDLP. In contrast, the LDLP's beta weight drops from .19 to .08 when all predictions enter the equation, a drop of nearly 60%. When only the AFQT, the ALQ, and the LDLP are used to predict the criterion, the LDLP and ALQ retain substantial beta weights. Considering these results, it appears that the LDLP scales predict some portion of the variance in the criterion scores that is also predicted by the ABLE scales and the Accelerated Promotions measure from the PFF.

Job-Related Subgroup Validity Generalization Analyses

Analyses were performed to determine if the validity results generalized across the job-related subgroups. Before these analyses could be done, cluster analyses were performed to identify the job-related subgroups.

Table 6.40

Beta Weights From Multiple Regressions Using AFQT as the Cognitive Ability Marker, Optimal Equations.

Predictors in Equation	Beta Weights								
	R_{CA}	ΔR	ALQ	LDLP	AFQT	DEP	DOM	WOR	PRO
AFQT	07				10				
AFQT, DEP, DOM, WOR	38				11	-09	17	26	
AFQT, DEP, DOM, WOR, ACCP	39				06	-08	16	24	13
ALQ, LDLP	28		20	19					
AFQT, ALQ, LDLP	28	21	20	19	02				
AFQT, DEP, DOM, WOR, ALQ, LDLP	41	03	16	08	05	-10	15	23	
AFQT, DEP, DOM, WOR, ACCP, ALQ, LDLP	41	02	15	08	02	-09	13	22	10

Note: n = 407. Decimals omitted. SE scores include SE-ALQ, PNC, and LDLP; R_{CA} = Multiple R, corrected for range restriction, adjusted for shrinkage; ΔR = change in R_{CA} with and without the ALQ in equation; ALQ = Army Leadership Questionnaire; AFQT = Armed Forces Qualification Test; AO = Assembling Objects. ABLE scale names abbreviated as follows: ADJ = Adjustment; COO = Cooperation; DEP = Dependability; DOM = Dominance; LOC = Locus of Control; PHY = Physical Condition; WOR = Work Orientation. PFF measures abbreviated as follows: REC = Recognition; DA = Disciplinary Action; PRO = Recommendation for Accelerated Promotions.

Cluster Analyses

Cluster analyses were performed to identify appropriate subgroups for the validity generalization analyses. The goal was to form subgroups based upon similar job task profiles. Two relevant measures were used: the Job Activities Questionnaire (JAQ) and one of the questions on the Biographical Information Form (BIF). On the JAQ, NCOs rated the importance and frequency of 13 task dimensions in performing their jobs. A composite variable--the product of importance and frequency--was computed for each of the 13 task dimensions. These 13 composite variables were used in the JAQ cluster analyses. On the BIF, NCOs indicated which of 10 jobs their subordinates held. An NCO could indicate any number of jobs (i.e., from 0 to 10). This resulted in a set of 10 clustering variables where each variable had a value of 0 (some subordinates held this job) or 1 (no subordinates held this job).

Several cluster analyses were attempted before the best cluster solution was chosen. Each preliminary analysis is briefly described below. Ward's method of clustering was used in each analysis.

In the first set of cluster analyses, individual NCOs were clustered. The proportion of variance retained by a reasonable number clusters (e.g., 20 or fewer clusters) was very small for both the JAQ profiles and the subordinate job profiles. The cubic clustering criterion value indicated that there were about 80 clusters of NCOs. (The data were orthogonalized on the clustering variables before computing the cubic clustering criterion.)

In subsequent analyses, clustering was done at the MOS level. All NCO's indicating membership in the same MOS (on their Background Information Form) were used to compute mean values on the clustering variables. This strategy was based on the assumption that NCOs within the same MOS performed similar tasks (and that the variability found within MOS in the previous cluster analyses was due to measurement error). Therefore, it was also assumed that validity coefficient values generalize within an MOS. The new goal of the clustering analyses was to form clusters of MOSs where the JAQ and subordinate job profiles of the various MOSs within a cluster were similar.

Cluster analyses were performed using the JAQ profiles or subordinate job profiles. The cluster analysis using the JAQ profiles resulted in a reasonable number of clusters. In the hierarchical clustering iterations, there was a discontinuity in the $\underline{R}^2$ values between 11 and 12 clusters. At 12 clusters, a substantial amount of variance--59 percent--was retained. There was a serious problem, however, with this 12-cluster solution: The MOS titles within a cluster appeared to represent very diverse MOSs, and MOSs with very similar titles were often in different clusters. Therefore, this clustering solution was not used.

The cluster analysis using the subordinate job profiles also formed a reasonable number of clusters. It was decided to retain 12 clusters. This decision was based on a discontinuity in the $\underline{R}^2$ values between 11 and 12 clusters and the substantial amount of variance--77 percent--retained. This cluster analysis, in contrast to the cluster analysis using the JAQ profiles, formed clusters with similar MOS titles. For example, one cluster contained all the medical MOSs. This

cluster analysis also retained more variance than the JAQ cluster analysis (77 percent vs. 59 percent). Therefore, the cluster analysis for the subordinate jobs was used.

Because validity generalization analyses with so many subgroups (i.e., 12 subgroups) would have had low statistical power and been impractical, the number of subgroups was further reduced. The 12 clusters from the subordinate jobs cluster analysis were rationally placed into three subgroups based on the MOS titles in the clusters. These three final subgroups were named Combat, Repair/Technical, and Administration. The MOS codes in each of the groups are shown below in the Validity Generalization Analyses section.

Validity Generalization Analyses

Analyses were performed to determine if the validity results generalized across the job-related subgroups identified in the cluster analyses. A meta-analysis was performed to answer this question for both the bivariate and multivariate validity coefficients.

The SAS IML regression program was used to determine the correlations between the overall criterion and several operational and experimental predictors across three job groups. The job groups, Group 1, Group 2, and Group 3 roughly correspond to Combat, Repair/Technical, and Administration, respectively. The three job groups and the MOSs they included were as follows:

Two Group 1 =00B, 11B, 11C, 11H, 11M, 12B, 12C, 12F, 13B, 13C, 13E, 13F, 13R, 13Z, 14R, 14S, 16T, 19B, 19K, 31C, 31D, 31F, 31M, 31R, 31U, 31W, 31Y, 54B, 71C

Group 2 = 24T, 25M, 27T, 35E, 35F, 35J, 35N, 35W, 45E, 45G, 45K, 51B, 51T, 62B, 62E, 62J, 62N, 63B, 63H, 63J, 63S, 63W, 63Z, 64O, 67N, 67R, 67S, 67T, 67U, 67V, 67Y, 68B, 68D, 68F, 68G, 68J, 68K, 68N, 68X, 68Y, 80K, 88H, 88L, 88M, 88N, 91B, 91C, 91G, 91P, 91S, 95B, 96B

Group 3 = 55B, 71L, 75B, 75D, 75E, 75F, 75Z, 77F, 92A, 92Y, 93B, 93P, 94B.

The predictors used in the analysis appear in Table 6.41, along with descriptive statistics for each of the three job groups.

The three group-specific correlations for each predictor were then analyzed to determine the degree to which their variation was due to sampling error. The procedure used was that described by Hunter, Schmidt, and Jackson (1982). Briefly, the variation in correlations is the result of two factors: variation in the population correlations and sampling error. Each sample correlation (r_i—the correlation from study I) is taken as the sum of the population correlation (ρ_i) and sampling error (e_i):

$$r_i = \rho_i + e_i$$

with the mean of the sampling error equal to zero and its variance equal to:

$$\sigma_{ei}^2 = \frac{(1 - \rho_i^2)^2}{N_i - 1}$$

With some simplifying assumptions, the estimated variance due to sampling error can be calculated as

$$\sigma_e^2 = \frac{(1 - \bar{r}^2)^2 K}{N}$$

where K is the number of studies and N is the total sample size across studies. (In the present analysis, K is the number of job groups—that is, the number of correlations available.) The variance across studies is given by

$$\sigma_r^2 = \sum \frac{[N_i(r_i - \bar{r})^2]}{\sum N_i}$$

where

$$\bar{r} = \frac{\sum [N_i r_i]}{\sum N_i} .$$

Thus, to correct the variation in observed correlations for sampling error, we work through the equations just given in reverse order: (a) calculate the weighted mean correlation across studies, (b) determine the observed variation in the correlations across studies, (c) calculate the estimated variation due to sampling error, and (d) subtract the variance due to sampling error from the observed variation,

$$est\ \sigma_\rho^2 = \sigma_r^2 - \sigma_e^2$$

This estimated variation in the population correlations can be tested for significance using the following statistic:

$$\chi^2_{K-1} = \frac{N}{(1 - \bar{r}^2)^2} \sigma_r^2 \quad .$$

Hunter et al. (1982, p. 47) reported that this test is very powerful; failure to find significant results are sure signs of no true variation in the population correlations, whereas significant results might be due instead to sampling error.

Table 6.41

Descriptive Statistics of the Selected Operational and Experimental Predictors.

	Job Group #1 (n = 199)		Job Group #2 (n = 141)		Job Group #3 (n = 25)	
	M	SD	M	SD	M	SD
AFQT	51.33	23.54	57.70	21.74	53.84	19.68
WK	14.54	2.94	15.47	2.85	15.16	2.90
AR	13.06	3.48	13.84	3.60	12.92	3.85
AO	15.19	3.56	15.87	3.37	14.16	3.13
ALQ	42.63	10.68	43.28	8.43	42.68	8.35
LPI	30.41	10.06	32.06	9.58	32.38	10.12
SE-ALQ	165.17	35.38	160.40	36.21	157.16	28.87
PNC	3.07	1.17	3.04	1.13	3.04	1.17
LDLP	3.75	1.15	3.86	1.13	4.16	1.17
ABLE - Dependability	51.39	5.66	51.99	6.32	53.68	6.17
ABLE - Dominance	46.56	5.20	45.86	5.74	47.92	4.00
ABLE - Work Orientation	68.48	7.88	69.04	7.55	70.97	8.28
Criterion-- Leadership Perf	150.27	23.29	153.03	22.40	150.35	19.12

The results of the meta-analysis are provided in Table 6.42 (no shrinkage correction) and Table 6.43 (shrinkage correction). Table 6.42 shows no significant variation across the unadjusted correlations. Five bivariate correlations and one multiple correlation yield non-zero standard deviations for the population correlations, but in no instance is the variation significant.

Table 6.43 presents similar results, except that now the chi-square tests are significant at $p < .05$ for the Assembling Objects measure and the composite of the AFQT and the LPI. These results were obtained because several of the adjusted correlations go to zero (job groups 1 and 2 for AO, job group 2 for the AFQT/LPI composite). Given the power of the chi-square test

(arguably lessened given the small total sample size of 265) and the likely overadjustment of the shrinkage formula, these results should not be given undue emphasis. Rather, the primary finding is that the variation across job groups for the selected predictors is minimal. Validation of these predictors for a general leadership criterion appear to generalize across types of NCO jobs, at least as we were able to identify those types in these data.

Fairness Analyses

We conducted differential prediction analyses to investigate differences between racial subgroups in the sample. We designed our analyses according to the Cleary (1968) definition of bias:

> A test is biased for members of a subgroup of the population if, in the prediction of a criterion for which the test was designed, consistent nonzero errors of prediction are made for members of the subgroup (p. 115).

Following this definition, fairness can be evaluated by comparing various predictor/criterion regression parameters associated with different subgroups. We engaged in the following steps to make this comparison:

(1) Compute least squares weighted composite scores (predicted performance scores) for individuals based on the regression of the predictor composites against the criterion variable. We created these scores for the following predictor combinations:

- AFQT, ALQ
- AFQT, LPI
- AFQT, LDLP
- AFQT, ALQ, LDLP

We used these composites because a) cognitive test scores are more likely to produce differences in subgroup scores than are temperament measure like the ABLE; b) it was unlikely that new measures would be used independently from ASVAB-based scores; and c) we wished to examine the most promising of the new measures.

(2) Regress the individual experimental predictors and least squares weighted composite scores against the criterion for each subgroup and the total group.
(3) Compute F-tests for regression slope homogeneity and for intercept differences.
(4) Graph the regression results.

As an example, Figure 6.3 shows the results of the White/Black differential prediction analyses for the AFQT/ALQ composite against the criterion. It provides a number of pieces of information:

Table 6.42

Subgroup Analyses of Regression Results, Not Adjusted for Shrinkage.

	Cluster 1	Cluster 2	Cluster 3	$\bar{r}$	var r	var e	var rho	sd rho	chi sq*	p values
N	199	141	25							
AFQT	0.16	0.08	0.48	0.1510	0.0094	0.0081	0.0013	0.0366	3.5950	0.1657
WK	0.16	0.01	0.10	0.0979	0.0051	0.0081	-0.0030	0.0000		
AR	0.15	0.13	0.57	0.1710	0.0118	0.0081	0.0037	0.0611	4.5682	0.1019
AO	0.04	0.08	0.54	0.0897	0.0153	0.0081	0.0072	0.0849	5.6649	0.0589
ALQ	0.30	0.24	0.12	0.2645	0.0023	0.0081	-0.0057	0.0000		
LPI	0.21	0.04	0.42	0.1587	0.0116	0.0081	0.0035	0.0591	4.4381	0.1087
SE--ALQ	0.02	0.17	0.08	0.0821	0.0051	0.0081	-0.0030	0.0000		
PNC	0.16	0.05	0.14	0.1161	0.0028	0.0081	-0.0053	0.0000		
LDLP	0.24	0.20	0.08	0.2136	0.0017	0.0081	-0.0064	0.0000		
ABLE--DEPN	0.05	0.01	0.18	0.0435	0.0017	0.0081	-0.0063	0.0000		
ABLE--DOMN	0.36	0.36	0.28	0.3545	0.0004	0.0081	-0.0077	0.0000		
ABLE--WORK	0.44	0.33	0.04	0.3701	0.0107	0.0081	0.0027	0.0518	5.2674	0.0718
AFQT+ALQ	0.31	0.24	0.48	0.2946	0.0036	0.0081	-0.0044	0.0000		
AFQT+LPI	0.23	0.09	0.57	0.1992	0.0145	0.0081	0.0065	0.0805	5.7552	0.0563
AFQT+LDLP	0.26	0.22	0.54	0.2637	0.0060	0.0081	-0.0021	0.0200		
AFQT+ALW+LDLP	0.36	0.30	0.55	0.3498	0.0038	0.0081	-0.0043	0.0000		

*Hunter, Schmidt, and Jackson (1982, p.47) *Notes:* Chi-square df = k-1 = 2. Correlations are corrected for range restriction.

Table 6.43
Subgroup Analyses of Regression Results, Adjusted for Shrinkage.

	Cluster 1	Cluster 2	Cluster 3	r	var r	var e	var rho	sd rho	chi sq*	p values
N	199	141	25							
AFQT	0.12	0.00	0.47	0.0976	0.0135	0.0081	0.0054	0.0734	5.0050	0.0819
WK	0.13	0.00	0.00	0.0709	0.0042	0.0081	-0.0039	0.0000		
AR	0.11	0.08	0.57	0.1299	0.0144	0.0081	0.0064	0.0799	5.4547	0.0654
AO	0.00	0.00	0.53	0.0363	0.0179	0.0081	0.0099	0.0993	6.5588	0.0377
ALQ	0.29	0.22	0.07	0.2479	0.0034	0.0081	-0.0046	0.0000		
LPI	0.18	0.00	0.41	0.1262	0.0132	0.0081	0.0052	0.0720	4.9930	0.0824
SE-ALQ	0.00	0.14	0.00	0.0541	0.0046	0.0081	-0.0034	0.0000		
PNC	0.13	0.00	0.10	0.0777	0.0039	0.0081	-0.0042	0.0000		
LDLP	0.22	0.17	0.00	0.1856	0.0031	0.0081	-0.0050	0.0000		
ABLE--DEPN	0.00	0.00	0.15	0.0103	0.0014	0.0081	-0.0066	0.0000		
ABLE--DOMN	0.35	0.34	0.26	0.3400	0.0005	0.0081	-0.0076	0.0000		
ABLE--WORK	0.43	0.32	0.00	0.3581	0.0122	0.0081	0.0041	0.0640	5.8410	0.0539
AFQT+ALQ	0.28	0.20	0.47	0.2621	0.0046	0.0081	-0.0034	0.0000		
AFQT+LPI	0.19	0.00	0.55	0.1413	0.0204	0.0081	0.0124	0.1113	70770	0.0205
AFQT+LDLP	0.23	0.17	0.52	0.2267	0.0071	0.0081	-0.0009	0.0000		
AFQT+ALQ+LDLP	0.32	0.25	0.53	0.3073	0.0048	0.0081	-0.0033	0.0000		

Hunter, Schmidt, and Jackson (1982, p. 47)
Notes: Chi-square df = k-1 = 2. Correlations are corrected for range restriction.

- The effect size for both the test score difference and the criterion score difference. The White predictor composite mean was .72 SD higher than the Black test score mean, and the White criterion score mean was .02 SD higher than the Black score mean (note that the criterion score means are nearly equal to one another).

- The p value for the level of significance of the difference between (a) subgroup slopes and (b) intercepts. Both p values are based on the Type I Sum of Squares (Cohen & Cohen, 1983, p. 145). Neither the slope nor the intercept differences are significant (at the .05 level) in the example shown in Figure 6.2.

- Predictor scores for key points in the graph.

  - Predicted performance scores [i.e., the subgroup intercept + (the subgroup slope*the predictor score)] for the same key points.
  - The standard error of the predicted performance score (Hays, 1963, p. 522, formula 15.22.3).
  - The value of the difference between subgroup predicted performance scores "Under" indicates that performance of the protected group would be underpredicted by the Non-Protected group line. "Over" indicates overprediction when the subgroup lines are compared.

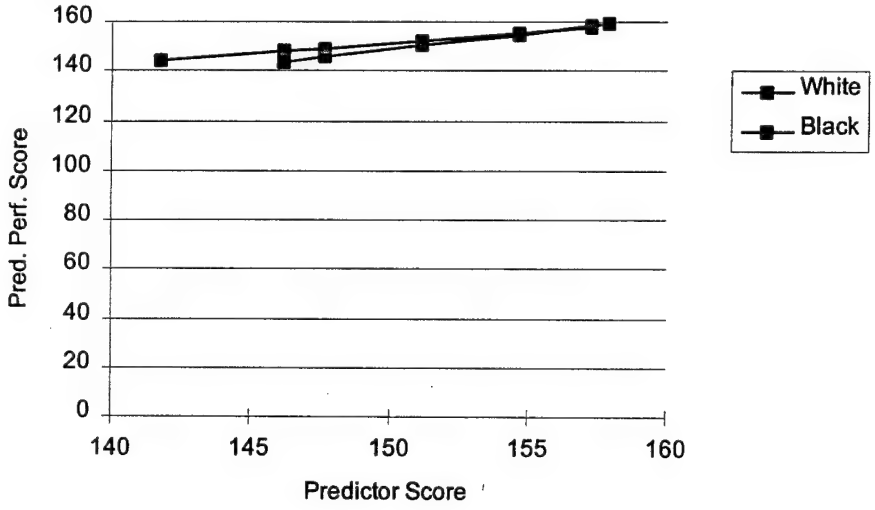
First, we conducted this analysis of differential prediction for the ALQ, the LPI, and the LDLP. Then we conducted these analyses for the predictor composites mentioned earlier. Our analyses focused on the racial subgroups of blacks and hispanics. There were too few females in our sample to provide useful differential prediction analyses by gender. The results of the differential prediction analyses are summarized in Table 6.44. The symbols in the cells of this table indicate significant ($p < .05$) slope and intercept differences and under- or overprediction. An "s" indicates significant slope differences; "I" indicates significant intercept differences; "u" and "o" indicate under- or overprediction (respectively). A "u" is assigned when there is underprediction at two or more of the following predictor scores: (a) the total group mean, (b) the total group mean plus one standard deviation, and (c) the total group mean minus one standard deviation. An "o" indicates overprediction at two or more of the three points. With to the Black/White comparisons, none of the comparisons yielded a significant difference. This is true for the individual predictors and the predictor composites.

With regard to the Hispanic/White comparisons, one out of the nine comparisons yielded a significant difference. Where the LDLP was used as the only predictor of the criterion, there was a significant ($p < .05$) slope difference leading to overprediction of Hispanic performance. When the AFQT and LDLP were combined to predict performance, the slopes differed at the $p < .10$ level. Finally, when both the ALQ and the LDLP were combined with the AFQT to predict performance, no differences were found.

Differential Prediction Analysis: Black/White Comparison
Composite: AFQT, ALQ
Criterion: Leadership Performance

Group	N	Test MN	Test SD	Criterion MN	Criterion SD	Slope	Intercept	R-Squared	R
Total	383	151.16	3.49	151.16	22.73	1.00	0.00	0.02	0.15
White	251	152.03	2.93	151.29	22.30	1.33	-50.77	0.03	0.17
Black	132	149.51	3.87	150.92	23.60	0.88	19.27	0.02	0.14
Effect Size		0.72		0.02					
P Value						0.52	0.34		

Predictor Score		Predicted Performance Score		Standard Error		Score Difference	
In relation to Mean Predictor Score for Total Group	Value	White	Black	White	Black	Value	Direction
-3 SD	141.77	NA	144.02	NA	4.55	NA	NA
-2 SD	146.17	143.50	147.90	3.10	2.68	-4.40	Under
-1 SD	147.67	145.48	149.22	2.49	2.25	-3.73	Under
Mean	151.16	150.13	152.29	1.45	2.21	-2.16	Under
+1 SD	154.66	154.77	155.36	1.86	3.38	-0.59	Under
+2 SD	157.25	158.21	157.64	2.83	4.55	0.57	Over
+ 3 SD	157.89	159.07	NA	3.10	NA	NA	NA



Notes: Effect Size—the difference between subgroup means in standard deviation units. Negative values indicate higher male (or, for black/white comparisons, white) means. Predicted Performance Score—the subgroup intercept + (the subgroup slope *the predictor score). Standard Error—the standard error of the predicted performance score (Hays, 1963, p. 522, formula 15.22.3). Score Difference—the value is the difference between subgroup predicted performance scores. For direction, the word "under" indicates that the performance of the protected group would be under predicted by the majority group line. The word "over" indicates overprediction.

Figure 6.3 White/Black differential prediction analysis.

Three main points can be drawn from the fairness analyses:

- There were no predictive differences between Blacks and Whites.
- A predictive difference between Hispanics and Whites was observed with the LDLP. This resulted in overprediction of Hispanic performance, which is generally not considered unfair test use (SIOP, 1987).
- The difference was alleviated when the full composite including the AFQT, the ALQ, and the LDLP was used to predict the criterion.

Table 6.44

Slope and Intercept Comparisons for Predictors and Predictor Composites Across Racial Subgroups.

Predictors	Black/White	Hispanic/White
ALQ	--	--
LPI	--	--
SE-ALQ	--	--
PNC	--	--
LDLP	--	s,o
AFQT, ALQ	--	--
AFQT, LPI	--	--
AFQT, LDLP	--	--
AFQT, ALQ, LDLP	--	--

CHAPTER 7
CONCLUSIONS

The objective of this research has been to develop and evaluate new measures that may improve the prediction of NCO performance. We developed five new measures. Our basic conclusions regarding each of these measures follow:

(1) The ALQ has promise for improving the prediction of NCO performance, and thus should be given serious consideration for a variety of operational uses. The ALQ has a moderate (about .20 -.24) bivariate relationship with the criterion, but it substantially improved prediction of the criterion over the AFQT and other measures of cognitive ability (about .15). In addition, the ALQ exhibited no predictive differences between Whites and Black or Hispanic subgroups. However, the ALQ added only minimally to the predictive power of the AFQT combined with the ABLE (about .01 or .02).

The minimal incremental validity over the AFQT/ABLE combination may lead one to conclude that the ALQ is not worth pursuing further. However, there are advantages of the ALQ beyond this small incremental validity. The ALQ is not likely to be affected by demand characteristics, it has good face validity, and it could be applied in various ways beyond prediction of NCO performance. We detail these points below.

Scores on the ALQ may not be strongly affected by demand characteristics. Although previous research has shown a strong relationship between the ABLE and performance (McHenry, Hough, Toquam, Hanson, & Ashworth, 1990), follow-up research has shown that the validity of the ABLE decreases when a longitudinal design is used, and that its score characteristics change depending on demand characteristics (Oppler, Peterson, & Russell, 1993). Job applicants may tend to respond to the ABLE in a socially desirable way, leading to a ceiling effect on various scales, reductions in the variance of the scores, and attenuation of the validity of the measure. However to the extent that the ALQ is valid because it measures job knowledge, social desirability might be less of a concern.

The ALQ has obviously greater face validity than the ABLE. Applicants or especially, candidates for promotion, are likely to question the relevance of the ABLE items, and in some cases, view the items as intrusive. On the other hand, the ALQ asks respondents to consider "real-life" problems encountered by NCOs and choose among options written by other NCOs. Thus, the ALQ would likely meet with fewer objections from respondents and decision makers, and might ultimately see more use than other, less face valid instruments.

Research conducted on the ABLE shows that those items most susceptible to change are those more closely related to experience in the Army (White & Moss, 1995). The ALQ unabashedly measures experience-rich items. Perhaps, if the experience-based items in the ABLE were not used, then the ABLE, in concert with the ALQ, would be useful for the NCO population.

The ALQ could be applied in a variety of ways beyond use in selection or promotion. Apart from its obvious potential for helping to select NCOs, the ALQ could be used for training and career development. The ALQ's content rich items are ideal for development of more realistic testing involving the use of video or computers.

Finally, we think that an important element in this research was the effort to develop a taxonomic framework for the ALQ. The identification of 8 distinct situational categories and 3 levels of urgency or priority of situations enabled the systematic development of item content and a more straight-forward approach to developing alternate forms. One avenue for further investigation might focus on separate analyses of category subscores on the ALQ, and/or the expansion of items measuring these categories.

(2) The LPI did not show promise for incrementing the prediction of performance, and thus should probably not be pursued as an alternative instrument. The LPI failed to predict the criterion by itself, and did not serve as a useful increment of NCO performance. It was a noble experiment, but shows little promise in its present form.

(3) The LDLP contributed to the prediction of NCO performance, and thus its usefulness in an operational setting should be evaluated. The LDLP had a moderate bivariate relationship with the criterion (about .20), it substantially improved prediction of the criterion over the AFQT and other measures of cognitive ability, and, while it exhibited some predictive differences between racial and ethnic groups, these differences would not result in unfair test use.

The other self-efficacy scores (SE-ALQ and PNC) did not relate to NCO performance, and thus should not be pursued as alternative predictors, with one possible caveat, as noted earlier. If the SE-ALQ and PNC measures were used in combination with the ALQ, some incremental use might be found. For example, if the scores on SE-ALQ were inconsistent with observed ALQ scores (too high or too low), we might expect poorer job performance. These kinds of investigations could be conducted using the existing database.

Like the ALQ, the LDLP added only minimally, or not at all, to the predictive power of the AFQT and the ABLE. However, this minimal incremental prediction should not be cause to discard the LDLP. The LDLP contains content on issues commonly faced by NCOs. Therefore, the instrument

has high face validity, and thus might receive higher acceptance among both respondents and decision-makers. Also, the LDLP is a relatively short instrument in terms of testing time, meaning that administration costs would be low.

One potential difficulty with the LDLP is that it may be affected by demand characteristics, especially if promotion hinged on the outcome of the test. Research must be done to evaluate this issue. Perhaps the rich and specific content of the problems may lead respondents to respond in a veridical manner despite the demand characteristics. Alternatively, it may be that framing the question differently, (e.g., so that it refers to past performance) may reduce response distortion. At any rate, we would not recommend use of the LDLP for selection purposes until this issue is addressed.

What do these findings appear to tell us about NCO performance? It must be kept firmly in mind that we were looking primarily at the leadership or supervisory aspects of NCO performance in our research. For example, our structured interview criterion ratings included dimensions like "Motivating Others," "Demonstrating Effort," "Planning and Providing for Training," and "Organizing, Coordinating, and Executing." The situational supervisory ratings were based on situations that NCOs face, including: "Routine Requests of Complaints," "Conduct Problems," "Shortage of Resources," "Neglect of Responsibilities," " Personal Problems," " Complaints about Command Actions," "Skills Deficits," and "Subordinate Recognition or Promotion" (see Table 2.6). Given this part of the criterion space, the usual cognitive predictors did not do well. The best predictors were the ALQ, a tacit knowledge measure that focuses on the areas just mentioned, and the ABLE scales of Work Orientation and Dominance. PFF measures that showed early recognition and recommendations for accelerated promotion were also related to better performance. It appears that assertive, hard-working soldiers that are recognized as having potential, and go on to effectively acquire the tacit knowledge involved in first-line supervision will become effective NCOs.

There are various limitations to the study findings. First, our findings are based on a concurrent validation design. The incumbents in our sample were all NCOs, and thus differ in a number of ways from soldiers eligible to become NCOs. For example, NCOs in our sample may have had varying amounts of knowledge gained through training not available to non-NCOs. This knowledge may have led to improved performance on both the predictors and the criterion, thus increasing the correlation between the two. We contend, however, that much of the knowledge and skill needed to perform well as an NCO is probably not learned in training, and it is precisely these aspects of NCO performance that the ALQ was meant to predict.

The study findings may not generalize to situations with greatly different demand characteristics. One such situation might be that where these instruments are used as a promotion device. In such a situation, different demand characteristics may cause respondents to respond in a socially desirable manner or to distort their responses in some other way. Such response distortion might well reduce the validity of these predictors, although, as we have pointed out above, we think the ALQ is largely immune to such effects. There are other

applications for these instruments where the demand characteristics would not differ greatly (e.g., training), and thus not affect the usefulness of these instruments.

The relationships between amount of experience, effectiveness of NCO performance, and achievement on the ALQ or LDLP were not examined in this investigation. Years of experience is unlikely to be of much use in a system for promoting soldiers to the NCO ranks because soldiers generally will be considered for promotion within constricted ranges of experience (the correlation between rank and years of experience was .75 in our sample). The important predictive variance is the expected NCO performance variance between soldiers within an experience cohort, and the ALQ or LDLP are exactly the kinds of measures to gauge that variance.

If possible, future research should evaluate the ALQ and the LDLP using a predictive validation design, although there is reason to believe the ALQ results should hold up, based on prior Project A and Career Force research (Rumsey et al., 1996). Certainly, experiments should be conducted to investigate effects on ALQ and LDLP scores when they are used in "promotional" or other "high-stakes" situations. We would expect relatively little effect on ALQ scores, but there could be substantial effects on LDLP scores. Further studies could also evaluate these instruments against criteria focused on various other aspects of NCO performance, rather than an overall criterion. An effort should also be made to evaluate the fairness of these instruments across gender lines-- such an evaluation was not possible in the current study due to the low number of female participants. Finally various issues regarding the relationship between experience, NCO performance, and achievement on the ALQ and LDLP should be pursued in the future.

As a final note, many of the ancillary products of this research project should prove useful. The conceptualization and development of the structured interview dimensions and criterion rating scales are highly job-relevant and could be useful for a number of applied purposes. The investigation of the three PFF measures showed that the Recognition measure was directly correlated with NCO performance, while the Advanced Promotion measure contributed significantly in multiple regressions predicting NCO performance. These findings should contribute to the Army's on-going research into the development of and methods for identifying soldiers with higher potential for performing well as an NCO.

REFERENCES

Bandura, A. (1982). Self-efficacy mechanism in human agency. American Psychologist, 37, 122-147.

Brown, W. (1910). Some experimental results in the correlation of mental abilities. British Journal of Psychology, 3, 296-322.

Bruce, M. M. & Learner, D. B. (1958). A supervisory practices test. Personnel Psychology, 11, 207-216.

Campbell, J. P., & Zook, L. M. (Eds.) (1990). Building and retaining the Career Force: New procedures for accessing and assigning Army enlisted personnel - Annual Report, 1990 Fiscal Year (ARI Research Note 952). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. ADA 252 675.

Campbell, J. P., & Zook, L. M. (Eds.) (1992). Building and retaining the Career Force: New procedures for accessing and assigning Army enlisted personnel - Annual Report, 1992 Fiscal Year (ARI Research Note). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Campbell, J. P., & Zook, L. M. (Eds.) (1994). Building and retaining the Career Force: New procedures for accessing and assigning Army enlisted personnel - Annual Report, 1992 Fiscal Year (ARI Research Note). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Cleary, T. A. (1968). Test bias: Prediction of grades of Negro and white students in intergrated colleges. Journal of Educational Measurement, 5, 115-124.

Cohen, J., & Cohen, P. (1983). Applied multiple regression/correlation analysis for the behavioral sciences (2nd ed.) Hillsdale, NJ: Lawrence Erlbaum Associates.

Department of Defense (1982). Profile of American youth: 1980 nationwide administration of the Armed Services Vocational Aptitude Battery. Washington, DC: Office of the Assistant Secretary of Defense (Manpower Reserve Affairs, and Logistics).

File, Q. W. (1945). The measurement of supervisory quality in industry. Journal of Applied Psychology, 69, 323-337.

File, Q. W. & Remmer, H. H. (1971). How supervise? (Manual, Rev. ed.). New York: Psychological Corporation.

Flanagan, J. C. (1954). The critical incident technique. Psychological Bulletin, 51, 327-358.

Guilford, J. P. (1954). Psychometric Methods. New York: McGraw-Hill.

Guion, R.M. & Gottier, R.F. (1966). Validity of personality measures in personnel selection. Personnel Psychology, 18, 135-164.

Gulliksen, H. & Tucker, L.R. (1961). A general procedure for obtaining paired comparisons from multiple rank orders. Psychometrika, 26, 173-183.

Hoyt, C. (1941). Test reliability estimated by analysis of variance. Psychometrika, *6*, 153-160.

Hays, W. L. (1963). Statistics (1st ed.). New York: Holt Reinhart & Winston.

Hunter, J. E., Schmidt, F. L., & Jackson, B. J. (1982). Meta-analysis: Cumulating research findings across studies. Beverly Hills, CA: Sage.

Latham, G. P. & Saari, L. M. (1984). Do people do what they say? Further studies on the situational interview. Journal of Applied Psychology, 69, 569-573.

Latham, G. P., Saari, L. M., Pursell, E. D., & Campion, M. A. (1980). The situational interview. Journal of Applied Psychology, 65, 422-427.

Lawley, D. (1943). A note on Karl Pearson's selection formulae. Royal Society of Edinburgh, Proceedings, Section A, 1943-1944, 62, 28-30.

Lee, C. & Bobko, P. (1994). Self-efficacy beliefs: Comparison of five measures. Journal of Applied Psychology, 79, 364-369.

Lord, F. M. & Novick, M. R. (1968). Statistical theories of mental test scores. Reading, MA: Addison-Wesley.

McHenry, J. J., Hough, L. M., Toquam, J. L., Hanson, M. A., & Ashworth, S. (1990). Project A validity results: The relationship between predictor and criterion domains. Personnel Psychology, 43, 335-354.

Motowidlo, S. J., Dunnette, M. D., & Carter, G. W. (1990). An alternative selection procedure: The low-fidelity simulation. Journal of Applied Psychology, 75, 640-647.

Motowidlo, S. J. & Tippins, N. (1993). Further studies of the low-fidelity simulation interview in the form of a situational inventory. Journal of Occupational and Organizational Psychology, 66, 337-344.

Mowry, H. W. (1957). A measure of supervisory quality. Journal of Applied Psychology, 41, 405-408.

Oppler, S.H., Peterson, N.G., & Russell, T.L. (1993). Basic LVI validation results. In J.P. Campbell & L. Zook (Eds.), Building and retaining the career forces: FY 1991 annual report. Alexandria, VA: U.S. Army Research Institute for the Behavioral Sciences.

Peterson, N. G., Smith, D., Hoffman, R. G., Pulakos, E. D., Reynolds, D., Potts, B. C., Oppler, S. H., & Whetzel, D. L. (1993). Expanding the concept of quality in personnel: Base period final report. Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Peterson, N. G., Russell, T. L., Hallam, G., Hough, L. M., Owens-Kurtz, C., Gialluca, K., & Kerwin, K. (1990). Analysis of the experimental predictor battery: LV Sample. In J. P. Campbell, & L. M. Zook (Eds.) Building and retaining the Career Force: New procedures for

accessing and assigning Army enlisted personnel - Annual Report, 1990 Fiscal Year (ARI Research Note 952). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. ADA 252 675.

Peterson, N. G., Hough, L. M., Dunnette, M. D., Rosse, R. L., Houston, J. S., Toquam, J. L., & Wing, H. (1990). Project A: Specification of the predictor doamin and developmet of new selection/classification tests. Personnel Psychology, 43, 247-276.

Pulakos, E. D. (1991). Behavioral performance measures. In J. W. Jones, B. D. Steffy, & D. W. Bray (Eds.), Applying psychology in business: The manager's handbook. New York, NY: Lexington Books.

Pulakos, E. D., & Borman, W. C. (Eds.) (1986). Development and field test report for the Army-wide rating scales and the rater orientation and training program (Technical Rep. 716). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Rosen, N. A. (1961). How supervise? -- 1943-1960. Personnel Psychology, 14, 78-99.

Rosenberg, S. & Kim, M. J. (1975). The method of sorting as a data-gathering procedure in multivariate research. Multivariate Behavioral Research, 10, 489-502.

Rozeboom, W. W. (1978). Estimation of cross-validated multiple correlation: A clarification. Psychological Bulletin, 85, 1348-1351.

Ruch, W. W., Stang, S. W., McKillip, R. H., & Dye, D. H. (1994). Employee Aptitude Survey Technical Manual. Los Angeles: Psychological Services, Inc.

Rumsey, M.E., Peterson, N.G., Oppler, S.H. & Campbell, J.P. (1996). What's Happened Since Project A: The Future Career Force. Journal of the Washington Academy of Sciences, 84, 94-110.

Schmidt, F. L. (1996). Statistical significance testing and cumulative knowledge in psychology: Implications for training of researchers. Psychological Methods, 1, 115-129.

Schratz, M. K., & Ree, M. J. (1989). Enlisted selection and classifcation: Advances in testing. Military personnel measurement: Testing, assignment, evaluation (p. 1-40). New York: Prager.

Shrout, P. E. & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. Psychological Bulletin, 86, 420-428.

Society for Industrial and Organizational Psychology (1987). Principles for the validation and use of personnel selection procedures (3rd ed.) College Park, MD: Author.

Spearman, C. (1910). Correlation calculated from faulty data. British Journal of Psychology, 3, 271-295.

Tenopyr, M. L. (1969). The comparative validity of selected validity scales relative to success in production management. Personnel Psychology, 22, 77-85.

Torgerson, W. S. (1958). Theory and Methods of Scaling. New York, NY: John Wiley and Sons, Inc.

Ward, J. H. (1963). Hierarchial grouping to optimize an objective function. Journal of the American Statistical Association, 58, 236-244.

Weekley, J. A. & Gier, J. A. (1987). Reliability and validity of the situational interview for a sales position. Journal of Applied Psychology, 72, 484-487.

White, L. A., & Moss, M. C. (1995 May). Factors influencing concurrent validaties of personality constructs. Paper presented at the annual meeting of the American Psychological Association, Washington, DC.

Zaccaro, S. J., Gilbert, J. A., Marks, M.A., Connelly, M. S., & Mumford, M. D. (in preparation). Social skills influencing leader performance.

Zaccaro, S. J., Gilbert, J. A., Thor, K. K., & Mumford, M. D. (1991). Leadership and social intelligence: Linking social perceptiveness and behavioral flexibility to leader effectiveness. Leadership Quarterly, 2, 317-331.

Appendix A
Definitions of the 29 NCO Problem Categories

ECQUIP -- Leadership Problem Categories

1. Emotional Health Problems

Subordinates suffer from emotional health problems that prevent them from performing duties or affect their work relationships.

2. Legal Problems

Subordinate is involved in intentional or misleading activities of a criminal/illegal nature (e.g., drugs, since drugs are cause for discharge).

3. Activity Planning

NCO is tasked with a planing for a non-military activity with which he/she has little or no experience.

4. Extreme Weather Conditions

NCO is faced with extreme weather conditions that somehow impact on work, exercises, etc.

5. Equipment Problems (New, Broken, or Missing)

NCO is faced with situation that involves new, broken, or missing equipment.

6. Poor Subordinate Appearance

NCO has a subordinate who is not maintainting his/her physical appearance according to military standards.

7. Poor Billet Appearance

NCO has a subordinate who is not maintaining his/her living area according to military standards.

8. Inadequate Resources

NCO is faced with shortage of time, equipment, personnel, etc. to accomplish mission (e.g., not enough equipment or people to accomplish mission).

9. Changes in Work Methods and Procedures

NCO is faced with situation where ways of completing tasks are different than he/she is used to.

10. Conflict With Chain of Command

NCO is faced with a situation where he/she disagrees with the policies, procedures, actions, or decisions of a superior (including decisions that affect subordinates or must be conveyed to subordinates).

11. Subordinate Having Difficulty Adapting

NCO has subordinate who is having problems adapting to or coping with changes or new situations.

12. Gender/Racial and Fairness Issues

NCO is faced with a situation where the focus is on gender/racial issues or differential treatment of subordinates.

13. Subordinate Unprepared for Mission

NCO has task or mission to accomplish with unprepared/unskilled subordinates (e.g., has people but they aren't ready to do this mission).

14. Poor NCO/Superior Performance

NCO displays poor judgment, or does not act or plan adequately.

15. Poor Subordinate Conditioning

NCO has subordinate(s) who are not able to keep up during or complete PT.

16. Subordinate Personal Problem(s)

NCO has subordinate experiencing personal problems (causes are external to job).

17. Staffing Decisions and Problems

NCO is faced with decisions about staffing, training slots, leave requests, etc.; includes situations where military and family demands are in conflict (e.g., holiday leave).

18. Supervisor-Peer Role Conflict

NCO promoted into position where he/she must supervise friend(s) and/or former peers.

19. Subordinate Promotion/Award Decisions

NCO must make decisions about who to promote, recognize, or designate for an award.

20. Alcohol-Related Problems

NCO has subordinate who has problem due to drunkenness or alcohol overuse/abuse.

21. Subordinate Financial Problems

NCO has subordinate who is having financial difficulties.

22. Disrespect/Refusal to Follow Orders

NCO has subordinate refuse to follow order or show disrespect for rank.

23. Subordinate Poor Attitude or Behavior

NCO is faced with a subordinate who behaves in undisciplined or inappropriate manner, including laziness and overuse of sick call.

24. Command Change

NCO is faced with situation involving a new superior.

25. Poor Subordinate or Unit Effort

NCO faces situation where individual(s) or unit is/are not performing up to standard.

26. Extra Duty Requirements -- Difficult/Demanding Hours

NCO is faced with situation where individual or unit must work extra duty or extended hours, especially when motivation is or has declined.

27. Tardiness

NCO is faced with situation where subordinate is late to formations, etc.

28. Bureaucratic Snafu (a.k.a. System Failure)

NCO is faced with a problem due to bureaucratic error, red tape, etc.

29. Subordinate Career Uncertainty

NCO is faced with situation where subordinate needs help with career decision or direction.

Appendix B
Job Performance Category Definitions

Job Performance Category Definitions

A. Demonstrating Technical Knowledge and Skill. Possesses sufficient technical knowledge to perform effectively in own specialty; keeps informed of the latest developments in field; demonstrates competency in performing various job tasks.

B. Communicating Orally. Speaks in a clear, organized, logical manner; listens attentively to others and responds in a suitable manner; keeps others informed as necessary and appropriate; presents competent, understandable, and organized briefings and presentations; tailors the level of presentations for the audience, using appropriate style and mode; effectively presents information, even when faced with hostility or confrontation; presents positions clearly and persuades others to accept good ideas; argues effectively for position when appropriate.

C. Writing. Prepares written materials that are organized, logical, accurate, and contain all relevant information; uses language that is at an appropriate level for the audience; uses correct grammar, punctuation, and spelling; produces written materials that require little or no editing.

D. Demonstrating Effort and Initiative. Persists with high effort in the completion of work; takes independent action when necessary; seeks out and willingly accepts responsibility, extra work, and challenging assignments; conducts own work carefully, completely, and accurately ; persists in carrying out difficult assignments and follows through on assignments and responsibilities and responsibilities; works long hours as necessary to complete work; meets deadlines and time constraints; maintains high performance standards; evaluates own strengths and skill deficiencies; seeks out and volunteers for training or job assignments that will improve skills.

E. Following Regulations, Policies, and Procedures. Conscientiously follows prescribed procedures in carrying out duties and assignments; adheres to applicable policies and regulations; consults appropriate manuals or regulations to ensure that proper procedures are followed; accepts others' authority and follows orders; observes the chain of command.

F. Demonstrating Integrity and Discipline. Maintains high ethical standards; behaves in an unwaveringly correct, moral, and ethical manner; provides and accurate accounting of information (i.e., does not intentionally distort information); demonstrates trustworthiness; controls self-indulgence.

G. Relating and Cooperating with Others. Treats others in a courteous, diplomatic, and tactful manner; demonstrates consideration for others and concern for their welfare and safety; provides help and assistance to others; is attuned to changes in others' behavior that may be indicative of problems; asks for input from others and listens to others' ideas; treats others in a fair, objective, and unbiased manner; works effectively

as a team member or a team leader; resolves completing interests or needs by constructive compromise with others.

H. Motivating Others. Recognizes, encourages, and rewards effective job performance; submits paperwork necessary for awards, disciplinary actions, and performance reviews; corrects unacceptable conduct; uses disciplinary actions constructively.

I. Planning and Providing for Training. Evaluates/identifies training needs; institutes formal or informal programs to address training needs; ensures that training opportunities are provided to those in need; develops others by providing work experiences to enhance their skills and abilities; guides and assists subordinates on technical matters; demonstrates work tasks or procedures.

J. Directing, Monitoring, and Supervising Work. Assigns and delegates tasks to subordinates and others equitably and for maximum efficiency and effectiveness; correctly evaluates subordinates' strengths and weaknesses so as to assign work properly; sets goals, targets, and criteria for work and assignments; ensures that assignments are clearly understood; specifies proper procedures and ensures they are followed; makes sure that needed equipment and supplies are available to conduct the work; monitors performance and provides constructive feedback to ensure that subordinates' work is accurate and complete.

K. Organizing, Coordinating, and Executing. Organizes and coordinates work to accomplish objectives. Takes all information into account when deciding on a source of action; identifies potential problems or sources of problems and develops sound solutions.

L. Demonstrating Responsiveness. Modifies behavior or plans as necessary to reach goals; is able to maintain effectiveness in varying environments with various tasks, responsibilities, or people; demonstrates flexibility and openness to change.

M. Representing. Willingly represents the Army at community and social functions; presents a positive and professional image of self and the Army even when off work; shows pride in being a soldier; maintains proper military appearance and bearing; acts as an effective role model, commanding respect from others.

Appendix C
ALQ Item Statistics:
NCO Pilot Test Sample

Explanation:

This appendix includes the following statistics on each item and score type:

- N: number of respondents
- Mean: mean score
- S.D.: standard deviation of the scores
- Part-whole correlation: correlation between the score for that item using that algorithm, and that score type (i.e., MOST (MST), LEAST (LST), TOTAL (TOT)) with the scale score for a given algorithm, and score type. These are presented for every algorithm and score type.

Reading the table:

Item scores for each type and algorithm are listed along the rows. They are correlated with the total scores. The labels for the total scores are listed along the columns. The correlation between the part score and the total for the scoring algorithm and type are listed along the diagonals.

Note:

To develop the scoring key for Algorithms 3 and 4, we examined the USASMA keying sample responses, and determined the "correct" response to each item as that response with a clearly greater frequency of responses than the others. In some cases, there was no clear correct response. Thus, for algorithms 3 and 4, we did not code one for those items. The result is that certain items have a score of zero for algorithms 3 and 4. Those items are:

- Form A: 9, 15, 28, and 48
- Form B: 8, 10, 24, 39, and 48

FORM A

Itm	T-y-p-e	N	Mean	S.D.	ALGORITHM-1			ALGORITHM-2			ALGORITHM-3			ALGORITHM-4		
					MST	LST	TOT	MST	LST	TOT	MST	LST	TOT	MST	LST	TOT
1	A-1 MST	126	12.98	27.31	.21	.18	.21	.18	.19	.20	.18	.16	.18	.20	.18	.20
	LST	126	21.23	34.20	.21	.37	.33	.23	.38	.34	.22	.35	.32	.19	.37	.32
	TOT	126	34.21	51.46	.25	.34	.33	.25	.35	.33	.24	.32	.31	.23	.34	.32
	A-2 MST	126	26.58	14.59	.18	.12	.16	.17	.12	.15	.19	.08	.14	.17	.12	.15
	LST	126	36.17	26.87	.19	.34	.30	.20	.35	.31	.20	.33	.30	.18	.33	.29
	TOT	126	62.75	34.28	.23	.32	.30	.23	.33	.31	.23	.29	.29	.21	.31	.29
	A-3 MST	126	.40	.49	.15	.13	.15	.14	.12	.14	.17	.09	.14	.14	.12	.14
	LST	126	.50	.50	.19	.33	.29	.19	.35	.30	.19	.32	.29	.17	.32	.28
	TOT	126	.90	.80	.21	.29	.27	.21	.29	.28	.23	.26	.27	.20	.28	.26
	A-4 MST	126	.29	.64	.19	.19	.21	.17	.20	.20	.18	.16	.19	.18	.18	.20
	LST	126	.40	.66	.22	.40	.35	.25	.39	.35	.23	.37	.34	.19	.40	.34
	TOT	126	.70	1.08	.25	.35	.33	.25	.35	.33	.25	.32	.32	.23	.35	.33
2	A-1 MST	126	38.32	36.17	.50	.43	.50	.49	.40	.49	.46	.39	.47	.47	.40	.47
	LST	126	39.35	32.24	.25	.34	.33	.23	.34	.32	.21	.32	.30	.24	.34	.32
	TOT	126	77.67	50.17	.52	.53	.57	.50	.51	.55	.47	.49	.53	.49	.51	.55
	A-2 MST	126	45.75	27.87	.46	.42	.47	.46	.39	.46	.43	.38	.45	.43	.39	.45
	LST	126	47.41	24.87	.20	.35	.31	.17	.36	.30	.15	.34	.28	.18	.35	.31
	TOT	126	93.16	36.83	.48	.55	.56	.47	.54	.55	.43	.52	.53	.45	.54	.55
	A-3 MST	126	.63	.48	.45	.42	.47	.44	.39	.45	.43	.39	.44	.42	.39	.44
	LST	126	.63	.48	.20	.31	.28	.17	.32	.27	.15	.29	.25	.18	.30	.27
	TOT	126	1.26	.66	.47	.53	.55	.45	.52	.53	.42	.50	.51	.45	.51	.53
	A-4 MST	126	.60	.55	.50	.43	.50	.49	.40	.48	.46	.39	.46	.47	.40	.47
	LST	126	.62	.52	.25	.33	.32	.22	.34	.31	.20	.32	.29	.23	.33	.32
	TOT	126	1.21	.78	.51	.52	.56	.49	.51	.54	.46	.49	.52	.48	.50	.54
3	A-1 MST	126	27.65	18.88	.11	.07	.09	.13	.09	.12	.09	.09	.10	.10	.06	.09
	LST	126	19.43	24.68	.11	.41	.30	.11	.42	.30	.13	.43	.33	.12	.43	.32
	TOT	126	47.08	34.80	.14	.33	.26	.14	.35	.28	.14	.36	.29	.14	.34	.28
	A-2 MST	126	36.40	12.00	.16	.16	.17	.18	.18	.20	.14	.18	.18	.15	.16	.17
	LST	126	30.29	15.13	.07	.34	.24	.06	.35	.24	.10	.37	.27	.09	.35	.26
	TOT	126	66.69	21.37	.14	.33	.26	.15	.35	.28	.15	.36	.29	.15	.34	.28
	A-3 MST	126	.87	.33	.09	.07	.08	.09	.10	.11	.08	.10	.10	.09	.07	.09
	LST	126	.70	.46	.08	.33	.24	.07	.35	.24	.11	.37	.28	.10	.35	.27
	TOT	126	1.57	.62	.10	.28	.22	.11	.31	.24	.12	.33	.26	.12	.30	.24
	A-4 MST	126	.79	.57	.07	.01	.04	.08	.04	.07	.05	.04	.05	.07	.01	.04
	LST	126	.59	.68	.12	.43	.32	.12	.44	.32	.15	.45	.35	.13	.45	.34
	TOT	126	1.38	.98	.13	.30	.25	.13	.33	.26	.13	.34	.27	.13	.32	.26

FORM A (Continued)

4	A-1	MST	126	30.86	34.93	.26	.18	.23	.24	.20	.24	.26	.17	.23	.26	.18	.23
		LST	126	34.05	33.66	.21	.42	.35	.22	.43	.36	.21	.40	.34	.17	.41	.34
		TOT	126	64.90	55.75	.29	.37	.36	.28	.38	.37	.29	.34	.35	.27	.36	.35
	A-2	MST	126	39.91	22.24	.23	.15	.20	.22	.17	.21	.25	.14	.21	.23	.15	.20
		LST	126	43.40	27.13	.18	.37	.31	.18	.38	.32	.18	.35	.30	.15	.36	.30
		TOT	126	83.32	38.50	.25	.34	.33	.25	.37	.34	.27	.33	.33	.24	.34	.33
	A-3	MST	126	.63	.48	.21	.13	.18	.20	.16	.19	.24	.13	.20	.21	.14	.19
		LST	126	.59	.49	.16	.33	.28	.17	.35	.29	.17	.32	.28	.14	.33	.27
		TOT	126	1.21	.75	.24	.30	.30	.24	.33	.32	.27	.29	.31	.23	.31	.30
	A-4	MST	126	.56	.62	.24	.16	.21	.22	.18	.22	.25	.15	.21	.25	.15	.21
		LST	126	.56	.56	.21	.41	.35	.22	.42	.36	.21	.38	.34	.18	.40	.33
		TOT	126	1.11	.96	.28	.34	.34	.27	.36	.35	.28	.32	.33	.26	.33	.33
5	A-1	MST	125	15.13	20.27	.23	.13	.19	.25	.10	.19	.28	.09	.20	.25	.12	.19
		LST	125	11.57	16.19	.41	.36	.41	.39	.35	.40	.42	.36	.42	.41	.35	.41
		TOT	125	26.70	28.36	.40	.29	.37	.40	.27	.36	.44	.27	.38	.41	.29	.37
	A-2	MST	125	29.74	16.49	.15	.12	.14	.18	.09	.14	.21	.09	.16	.17	.13	.16
		LST	125	27.73	10.99	.39	.39	.42	.38	.40	.43	.39	.38	.43	.37	.37	.41
		TOT	125	57.47	20.48	.33	.30	.34	.35	.29	.34	.38	.27	.35	.34	.30	.35
	A-3	MST	125	.47	.50	.17	.14	.17	.20	.11	.16	.23	.11	.18	.18	.15	.18
		LST	125	.25	.43	.17	.15	.17	.14	.15	.16	.16	.20	.20	.17	.16	.18
		TOT	125	.72	.66	.24	.20	.24	.24	.18	.23	.27	.21	.26	.25	.21	.25
	A-4	MST	125	.43	.57	.16	.09	.13	.18	.06	.13	.21	.06	.14	.18	.09	.15
		LST	125	.18	.53	.26	.21	.25	.24	.20	.24	.27	.24	.28	.27	.22	.26
		TOT	125	.62	.83	.28	.19	.25	.28	.17	.24	.32	.19	.27	.30	.20	.27
6	A-1	MST	126	16.61	10.78	.20	.10	.15	.18	.10	.15	.17	.07	.13	.15	.06	.11
		LST	126	13.40	25.81	.26	.35	.34	.23	.34	.32	.26	.34	.34	.28	.34	.35
		TOT	126	30.02	27.75	.32	.36	.37	.29	.36	.36	.31	.34	.36	.32	.34	.36
	A-2	MST	126	27.69	5.75	.13	.14	.15	.14	.13	.14	.16	.11	.15	.12	.14	.14
		LST	126	29.33	16.12	.26	.34	.33	.22	.35	.32	.26	.33	.33	.27	.32	.33
		TOT	126	57.02	17.96	.27	.35	.34	.25	.35	.33	.28	.33	.34	.28	.33	.34
	A-3	MST	126	.79	.41	.16	.15	.17	.15	.14	.16	.16	.12	.15	.12	.13	.14
		LST	126	.35	.48	.21	.29	.28	.18	.29	.26	.21	.29	.28	.24	.29	.30
		TOT	126	1.13	.62	.27	.32	.32	.24	.31	.31	.27	.31	.32	.27	.31	.32
	A-4	MST	126	.78	.43	.17	.14	.17	.17	.14	.16	.18	.11	.15	.14	.12	.14
		LST	126	.17	.70	.27	.35	.34	.25	.33	.32	.27	.33	.34	.29	.35	.35
		TOT	126	.95	.85	.31	.36	.37	.29	.34	.35	.32	.33	.36	.31	.35	.36

FORM A (Continued)

7	A-1	MST	126	31.32	31.42	.29	.22	.27	.27	.23	.27	.29	.19	.26	.30	.19	.26
		LST	126	27.93	38.66	.29	.45	.41	.25	.46	.40	.26	.40	.37	.28	.41	.38
		TOT	126	59.25	57.68	.35	.42	.43	.32	.43	.41	.33	.37	.39	.35	.38	.40
	A-2	MST	126	40.33	19.25	.39	.32	.38	.38	.33	.39	.39	.28	.36	.40	.29	.37
		LST	126	36.34	27.13	.20	.32	.29	.16	.35	.29	.19	.31	.28	.21	.28	.28
		TOT	126	76.67	36.74	.36	.41	.42	.32	.43	.42	.34	.38	.40	.36	.36	.40
	A-3	MST	126	.83	.38	.38	.30	.37	.37	.30	.36	.39	.27	.35	.39	.28	.36
		LST	126	.47	.50	.18	.27	.25	.13	.31	.25	.17	.28	.25	.19	.24	.24
		TOT	126	1.29	.68	.34	.37	.39	.31	.40	.39	.34	.35	.38	.36	.33	.38
	A-4	MST	126	.76	.56	.27	.19	.24	.25	.19	.24	.27	.16	.23	.28	.17	.24
		LST	126	.32	.72	.28	.44	.40	.24	.45	.39	.25	.40	.36	.27	.40	.38
		TOT	126	1.08	1.05	.33	.40	.40	.30	.41	.39	.32	.36	.37	.34	.36	.38
8	A-1	MST	125	33.64	33.66	.33	.15	.25	.34	.13	.25	.35	.11	.24	.33	.15	.25
		LST	125	36.34	40.69	.53	.59	.61	.54	.59	.62	.50	.57	.59	.52	.60	.62
		TOT	125	69.98	60.22	.55	.48	.55	.56	.47	.56	.53	.45	.54	.54	.49	.56
	A-2	MST	125	40.79	23.81	.29	.11	.21	.31	.10	.22	.34	.08	.21	.30	.12	.22
		LST	125	45.10	27.22	.48	.51	.54	.47	.52	.54	.43	.51	.52	.48	.52	.55
		TOT	125	85.90	39.65	.50	.42	.49	.51	.42	.50	.50	.40	.49	.51	.43	.51
	A-3	MST	125	.54	.50	.20	.04	.12	.22	.03	.13	.26	.01	.13	.22	.05	.14
		LST	125	.65	.48	.46	.48	.51	.44	.49	.51	.41	.48	.49	.46	.48	.52
		TOT	125	1.18	.73	.43	.34	.41	.44	.34	.42	.44	.32	.41	.45	.35	.43
	A-4	MST	125	.49	.59	.25	.07	.17	.27	.07	.17	.29	.04	.17	.26	.08	.17
		LST	125	.58	.62	.54	.59	.62	.55	.58	.62	.50	.57	.59	.53	.60	.62
		TOT	125	1.06	.96	.50	.43	.50	.52	.42	.51	.50	.40	.49	.50	.44	.51
9	A-1	MST	126	10.69	18.16	.14	.03	.09	.18	-0.00	.09	.10	.01	.05	.10	.00	.05
		LST	126	-2.68	21.73	-0.11	-0.04	-0.08	-0.09	-0.11	-0.11	-0.11	-0.10	-0.12	-0.12	-0.08	-0.11
		TOT	126	8.01	32.59	.01	-0.01	-0.00	.04	-0.08	-0.02	-0.02	-0.06	-0.05	-0.02	-0.05	-0.05
	A-2	MST	126	29.83	18.62	.12	.00	.06	.16	-0.04	.06	.07	-0.03	.02	.08	-0.03	.02
		LST	126	20.26	4.39	.06	.13	.11	.06	.15	.11	.09	.12	.12	.09	.12	.12
		TOT	126	50.09	18.59	.13	.03	.08	.17	-0.01	.08	.10	.00	.05	.10	.00	.05
	A-3	MST	126	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		LST	126	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		TOT	126	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
	A-4	MST	126	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		LST	126	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		TOT	126	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00

FORM A (Continued)

10	A-1	MST	126	26.60	40.01	.41	.28	.37	.43	.26	.37	.36	.28	.35	.35	.29	.35
		LST	126	46.94	44.53	.29	.46	.42	.30	.45	.42	.27	.44	.40	.27	.46	.41
		TOT	126	73.54	69.29	.42	.46	.48	.44	.44	.48	.38	.45	.46	.37	.46	.46
	A-2	MST	126	35.85	29.87	.31	.20	.27	.33	.19	.28	.28	.23	.28	.26	.21	.26
		LST	126	57.91	34.28	.32	.50	.46	.34	.50	.46	.30	.50	.45	.30	.50	.46
		TOT	126	93.76	48.57	.42	.48	.49	.44	.47	.50	.38	.49	.49	.37	.49	.48
	A-3	MST	126	.44	.50	.29	.19	.26	.32	.19	.27	.27	.23	.27	.25	.21	.25
		LST	126	.71	.46	.33	.51	.46	.34	.50	.47	.30	.50	.45	.31	.51	.46
		TOT	126	1.15	.71	.41	.46	.47	.44	.45	.49	.38	.48	.48	.37	.47	.47
	A-4	MST	126	.38	.60	.40	.27	.36	.42	.26	.36	.36	.28	.35	.34	.28	.34
		LST	126	.63	.61	.27	.44	.40	.27	.43	.39	.25	.42	.38	.25	.44	.39
		TOT	126	1.02	.98	.41	.44	.46	.43	.43	.47	.37	.44	.45	.36	.45	.45
11	A-1	MST	125	25.70	27.68	.24	.19	.23	.24	.17	.22	.26	.17	.23	.24	.18	.22
		LST	125	22.62	28.42	.11	.33	.25	.14	.34	.27	.11	.31	.24	.10	.31	.24
		TOT	125	48.33	41.11	.24	.36	.33	.25	.35	.33	.25	.33	.32	.23	.33	.32
	A-2	MST	125	35.06	18.03	.25	.17	.22	.25	.13	.21	.27	.13	.22	.24	.16	.21
		LST	125	34.29	24.11	.08	.25	.19	.10	.27	.21	.07	.25	.19	.06	.23	.17
		TOT	125	69.34	27.88	.23	.32	.30	.25	.32	.32	.23	.30	.30	.21	.30	.29
	A-3	MST	125	.54	.50	.22	.14	.19	.22	.11	.17	.25	.12	.19	.22	.14	.19
		LST	125	.46	.50	.04	.19	.14	.07	.22	.16	.03	.20	.13	.03	.17	.12
		TOT	125	1.00	.62	.21	.26	.26	.24	.26	.27	.22	.25	.26	.20	.25	.25
	A-4	MST	125	.49	.59	.22	.17	.21	.22	.15	.20	.25	.16	.22	.22	.17	.21
		LST	125	.44	.54	.09	.26	.20	.12	.28	.22	.09	.26	.20	.09	.25	.20
		TOT	125	.93	.78	.23	.31	.30	.24	.31	.30	.25	.30	.30	.23	.30	.29
12	A-1	MST	126	28.23	21.23	.30	.16	.24	.27	.19	.25	.29	.18	.25	.31	.15	.24
		LST	126	25.33	22.55	.29	.50	.44	.26	.48	.41	.33	.46	.44	.29	.48	.44
		TOT	126	53.56	32.72	.39	.45	.46	.35	.45	.45	.41	.43	.47	.40	.43	.45
	A-2	MST	126	33.71	11.21	.28	.14	.22	.24	.17	.22	.27	.16	.23	.28	.13	.22
		LST	126	33.56	14.35	.29	.48	.43	.26	.47	.41	.31	.45	.43	.29	.46	.42
		TOT	126	67.28	18.72	.39	.46	.46	.34	.46	.45	.40	.44	.47	.39	.43	.45
	A-3	MST	126	.83	.38	.22	.10	.17	.18	.13	.17	.21	.12	.18	.22	.09	.16
		LST	126	.74	.44	.26	.40	.36	.23	.39	.34	.29	.38	.37	.27	.39	.37
		TOT	126	1.56	.60	.33	.36	.37	.28	.37	.36	.34	.36	.39	.34	.34	.37
	A-4	MST	126	.76	.56	.31	.16	.25	.28	.19	.25	.29	.18	.25	.31	.15	.24
		LST	126	.67	.60	.27	.46	.41	.24	.44	.38	.31	.43	.42	.28	.45	.41
		TOT	126	1.43	.88	.38	.42	.44	.34	.43	.42	.40	.41	.45	.39	.40	.44

FORM A (Continued)

13	A-1	MST	125	18.17	32.89	.21	.15	.19	.23	.12	.19	.17	.09	.14	.15	.12	.15
		LST	126	4.33	30.96	-0.14	.02	-0.05	-0.09	-0.04	-0.07	-0.14	-0.01	-0.07	-0.18	.03	-0.07
		TOT	125	22.36	53.27	.05	.12	.10	.10	.06	.08	.03	.06	.05	-0.01	.10	.06
	A-2	MST	125	33.18	25.43	.20	.16	.20	.23	.12	.19	.17	.10	.14	.14	.14	.15
		LST	126	22.06	12.99	-0.14	-0.02	-0.08	-0.09	-0.08	-0.09	-0.14	-0.05	-0.10	-0.18	-0.01	-0.09
		TOT	125	55.21	31.34	.11	.13	.13	.15	.07	.12	.08	.06	.08	.04	.11	.09
	A-3	MST	125	.45	.50	.19	.16	.19	.22	.12	.19	.17	.10	.14	.13	.14	.15
		LST	126	.25	.43	-0.14	.01	-0.06	-0.10	-0.04	-0.08	-0.15	.02	-0.06	-0.19	.04	-0.07
		TOT	125	.70	.70	.04	.11	.09	.09	.05	.08	.02	.07	.05	-0.03	.12	.06
	A-4	MST	125	.37	.63	.20	.16	.19	.22	.12	.18	.17	.10	.15	.16	.13	.16
		LST	126	.07	.64	-0.16	.05	-0.04	-0.11	-0.01	-0.06	-0.16	.04	-0.05	-0.20	.07	-0.05
		TOT	125	.44	1.04	.02	.13	.09	.06	.07	.07	-0.00	.09	.05	-0.03	.12	.06
14	A-1	MST	125	13.98	24.84	.34	.26	.32	.30	.25	.30	.32	.25	.31	.37	.26	.34
		LST	125	7.83	21.10	.03	.09	.07	.00	.05	.03	-0.00	.09	.05	.03	.12	.08
		TOT	125	21.82	35.28	.25	.24	.27	.22	.20	.23	.22	.23	.25	.28	.25	.29
	A-2	MST	125	29.50	16.19	.35	.29	.35	.32	.27	.32	.31	.27	.32	.37	.28	.35
		LST	125	23.57	8.06	.04	.08	.07	.04	.03	.04	.04	.08	.07	.05	.11	.09
		TOT	125	53.07	19.36	.31	.28	.32	.28	.24	.28	.28	.26	.30	.33	.28	.33
	A-3	MST	125	.36	.48	.23	.14	.19	.16	.15	.17	.26	.15	.22	.30	.15	.24
		LST	125	.50	.50	.01	.06	.04	.01	.02	.01	-0.01	.06	.04	.00	.09	.05
		TOT	125	.86	.66	.17	.15	.17	.12	.13	.13	.18	.15	.18	.22	.17	.21
	A-4	MST	125	.18	.71	.37	.23	.32	.32	.23	.30	.36	.22	.31	.42	.23	.34
		LST	125	.44	.61	.06	.08	.08	.06	.03	.05	.06	.06	.07	.07	.11	.10
		TOT	125	.62	.97	.31	.22	.28	.27	.18	.25	.30	.20	.27	.35	.23	.31
15	A-1	MST	125	4.45	21.50	.25	.24	.27	.24	.20	.24	.19	.18	.20	.19	.22	.23
		LST	125	6.85	24.73	.16	.20	.20	.16	.15	.17	.15	.10	.13	.11	.14	.14
		TOT	125	11.30	37.73	.25	.27	.28	.25	.21	.25	.21	.17	.20	.18	.22	.22
	A-2	MST	125	23.28	13.23	.22	.20	.23	.21	.16	.20	.15	.15	.17	.15	.18	.19
		LST	125	25.29	13.51	.18	.28	.26	.18	.24	.23	.16	.18	.19	.13	.22	.20
		TOT	125	48.57	21.65	.24	.30	.30	.24	.25	.27	.19	.21	.22	.18	.25	.24
	A-3	MST	125	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		LST	125	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		TOT	125	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
	A-4	MST	125	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		LST	125	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		TOT	125	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00

FORM A (Continued)

16	A-1	MST	126	26.91	27.45	.40	.25	.34	.35	.25	.33	.39	.24	.34	.41	.24	.34
		LST	126	35.30	45.95	.37	.47	.47	.33	.50	.46	.36	.47	.46	.37	.46	.46
		TOT	126	62.21	58.53	.48	.49	.53	.43	.51	.52	.47	.48	.52	.48	.47	.52
	A-2	MST	126	33.82	16.87	.39	.25	.34	.35	.26	.33	.38	.24	.33	.40	.23	.33
		LST	126	44.56	34.69	.35	.43	.43	.31	.46	.43	.34	.43	.43	.36	.42	.43
		TOT	126	78.37	40.89	.46	.47	.50	.41	.49	.50	.44	.47	.50	.47	.45	.50
	A-3	MST	126	.75	.43	.35	.25	.32	.30	.25	.30	.33	.23	.31	.36	.22	.31
		LST	126	.53	.50	.34	.41	.41	.30	.44	.41	.33	.42	.42	.35	.40	.42
		TOT	126	1.29	.69	.47	.45	.50	.40	.47	.48	.45	.45	.49	.48	.43	.50
	A-4	MST	126	.71	.53	.39	.25	.34	.34	.25	.32	.38	.24	.33	.40	.24	.34
		LST	126	.37	.74	.38	.50	.48	.34	.53	.48	.37	.49	.48	.37	.48	.48
		TOT	126	1.09	.99	.49	.51	.54	.43	.53	.53	.48	.49	.54	.49	.49	.54
17	A-1	MST	126	25.19	27.15	.26	.26	.28	.28	.26	.29	.23	.23	.25	.20	.25	.25
		LST	126	20.42	31.30	.27	.40	.37	.29	.34	.35	.22	.34	.32	.21	.38	.34
		TOT	126	45.61	45.82	.33	.43	.42	.36	.39	.41	.29	.37	.37	.26	.41	.38
	A-2	MST	126	35.44	19.80	.19	.22	.23	.21	.21	.23	.17	.19	.20	.13	.21	.19
		LST	126	31.98	18.46	.21	.33	.30	.23	.29	.28	.17	.28	.25	.16	.31	.27
		TOT	126	67.42	29.14	.26	.36	.34	.29	.33	.34	.22	.31	.30	.19	.34	.30
	A-3	MST	126	.52	.50	.12	.19	.17	.14	.19	.18	.11	.16	.15	.07	.18	.14
		LST	126	.51	.50	.17	.28	.25	.19	.23	.23	.13	.23	.20	.12	.25	.21
		TOT	126	1.02	.75	.20	.31	.28	.22	.28	.28	.15	.26	.23	.13	.29	.24
	A-4	MST	126	.49	.55	.24	.27	.28	.26	.26	.28	.22	.23	.25	.18	.25	.24
		LST	126	.45	.60	.28	.43	.39	.30	.38	.38	.24	.37	.34	.23	.41	.36
		TOT	126	.94	.91	.33	.44	.43	.35	.40	.42	.29	.39	.38	.26	.42	.39
18	A-1	MST	126	16.42	22.13	.20	.19	.21	.19	.18	.20	.18	.14	.18	.20	.17	.20
		LST	126	7.75	20.01	.12	.25	.21	.13	.25	.21	.10	.28	.22	.10	.26	.21
		TOT	126	24.17	29.29	.23	.31	.30	.23	.30	.29	.21	.30	.28	.22	.31	.30
	A-2	MST	126	29.12	14.41	.28	.22	.27	.29	.21	.27	.25	.19	.24	.26	.21	.25
		LST	126	24.63	12.17	.13	.28	.23	.15	.28	.24	.16	.29	.25	.13	.30	.25
		TOT	126	53.75	19.33	.29	.34	.35	.31	.34	.35	.29	.32	.34	.27	.35	.35
	A-3	MST	126	.45	.50	.25	.29	.30	.23	.27	.28	.26	.24	.27	.26	.27	.29
		LST	126	.25	.43	.03	.15	.11	.06	.16	.13	.05	.19	.14	.02	.17	.12
		TOT	126	.70	.62	.23	.34	.31	.23	.33	.31	.24	.32	.31	.22	.33	.31
	A-4	MST	126	.40	.59	.18	.20	.20	.15	.18	.18	.16	.15	.17	.18	.17	.19
		LST	126	.20	.50	.15	.31	.26	.17	.31	.26	.16	.33	.28	.14	.33	.27
		TOT	126	.60	.79	.23	.34	.32	.22	.33	.31	.22	.32	.30	.23	.34	.32

FORM A (Continued)

19	A-1	MST	126	38.58	40.91	.28	.20	.25	.26	.22	.26	.28	.20	.26	.28	.20	.26
		LST	126	16.41	33.24	.15	.23	.21	.19	.23	.23	.17	.21	.21	.14	.22	.20
		TOT	126	54.99	57.87	.28	.28	.30	.29	.28	.31	.30	.26	.30	.27	.27	.30
	A-2	MST	126	47.25	27.64	.30	.24	.29	.30	.26	.30	.32	.24	.30	.30	.25	.29
		LST	126	27.37	22.25	.08	.15	.13	.12	.15	.15	.10	.14	.14	.08	.14	.12
		TOT	126	74.63	36.27	.28	.28	.30	.30	.29	.32	.31	.27	.32	.27	.27	.30
	A-3	MST	126	.66	.47	.31	.26	.31	.31	.28	.32	.33	.26	.32	.31	.27	.31
		LST	126	.31	.46	.04	.10	.08	.07	.10	.10	.06	.10	.09	.03	.09	.07
		TOT	126	.97	.65	.25	.26	.28	.28	.27	.30	.28	.26	.30	.25	.26	.28
	A-4	MST	126	.58	.63	.27	.19	.24	.26	.20	.25	.28	.19	.25	.27	.19	.25
		LST	126	.25	.55	.13	.21	.19	.17	.20	.21	.15	.19	.19	.12	.20	.18
		TOT	126	.83	.92	.27	.26	.28	.28	.26	.30	.28	.24	.29	.26	.25	.28
20	A-1	MST	125	30.11	39.91	.33	.26	.32	.33	.26	.32	.32	.26	.32	.32	.27	.32
		LST	124	17.09	28.80	.35	.38	.39	.32	.37	.38	.32	.38	.39	.34	.37	.39
		TOT	124	47.84	53.37	.43	.41	.45	.42	.39	.44	.40	.41	.44	.41	.40	.45
	A-2	MST	125	41.85	31.28	.33	.31	.34	.34	.30	.35	.32	.30	.34	.31	.31	.34
		LST	124	27.35	19.96	.20	.27	.26	.15	.27	.24	.18	.29	.27	.23	.27	.28
		TOT	124	69.51	38.84	.37	.39	.41	.35	.38	.40	.35	.39	.41	.37	.39	.42
	A-3	MST	125	.55	.50	.33	.32	.35	.35	.31	.36	.32	.31	.35	.32	.32	.35
		LST	124	.37	.48	.16	.23	.21	.11	.23	.19	.14	.26	.23	.19	.23	.23
		TOT	124	.93	.71	.34	.38	.39	.31	.38	.38	.32	.39	.40	.35	.38	.40
	A-4	MST	125	.50	.60	.33	.26	.31	.33	.26	.32	.32	.26	.32	.32	.26	.32
		LST	124	.32	.56	.30	.35	.35	.27	.34	.34	.28	.35	.35	.31	.34	.36
		TOT	124	.83	.88	.41	.40	.44	.39	.39	.43	.39	.41	.44	.41	.40	.45
21	A-1	MST	126	26.89	30.44	.34	.14	.25	.32	.15	.25	.30	.13	.23	.30	.12	.22
		LST	126	34.87	28.97	.27	.32	.32	.26	.33	.33	.26	.28	.30	.26	.30	.31
		TOT	126	61.75	46.88	.38	.29	.36	.37	.30	.36	.35	.26	.33	.35	.27	.34
	A-2	MST	126	33.64	19.35	.29	.13	.22	.29	.13	.22	.27	.11	.20	.25	.12	.19
		LST	126	41.28	21.31	.32	.32	.34	.30	.33	.35	.30	.28	.32	.31	.30	.33
		TOT	126	74.92	30.68	.40	.31	.38	.39	.31	.38	.38	.26	.35	.37	.28	.35
	A-3	MST	126	.72	.45	.27	.09	.18	.28	.09	.19	.25	.07	.17	.24	.08	.16
		LST	126	.83	.38	.22	.34	.31	.21	.35	.31	.21	.33	.30	.20	.33	.30
		TOT	126	1.55	.65	.32	.26	.31	.31	.26	.31	.30	.24	.29	.29	.25	.29
	A-4	MST	126	.61	.68	.35	.14	.25	.33	.15	.25	.31	.13	.24	.32	.13	.23
		LST	126	.75	.59	.19	.35	.30	.19	.36	.30	.18	.33	.29	.16	.35	.29
		TOT	126	1.36	1.03	.34	.29	.34	.33	.30	.34	.31	.27	.32	.30	.28	.32

FORM A (Continued)

22	A-1	MST	126	24.84	34.03	.27	.17	.23	.24	.19	.24	.27	.18	.24	.29	.16	.24
		LST	125	20.52	26.27	.28	.44	.40	.27	.42	.38	.31	.42	.40	.30	.45	.42
		TOT	125	45.87	49.49	.34	.35	.38	.32	.35	.37	.35	.34	.38	.37	.35	.39
	A-2	MST	126	34.85	19.58	.28	.16	.23	.25	.18	.23	.28	.16	.23	.31	.15	.24
		LST	125	28.60	12.26	.22	.38	.33	.22	.37	.33	.26	.36	.35	.25	.38	.36
		TOT	125	63.70	25.99	.32	.30	.33	.30	.31	.33	.34	.29	.34	.36	.29	.35
	A-3	MST	126	.74	.44	.27	.16	.22	.24	.18	.22	.27	.16	.23	.30	.15	.24
		LST	125	.86	.34	.30	.46	.42	.29	.43	.40	.32	.43	.42	.32	.47	.44
		TOT	125	1.61	.64	.35	.35	.38	.33	.35	.37	.36	.34	.38	.38	.35	.40
	A-4	MST	126	.48	.88	.27	.16	.22	.24	.18	.22	.27	.16	.23	.30	.15	.24
		LST	125	.73	.69	.30	.46	.42	.29	.43	.40	.32	.43	.42	.32	.47	.44
		TOT	125	1.22	1.29	.35	.35	.38	.33	.35	.37	.36	.34	.38	.38	.35	.40
23	A-1	MST	125	15.86	33.11	.25	.12	.19	.25	.08	.17	.24	.09	.18	.23	.11	.18
		LST	125	14.67	21.59	.23	.32	.31	.23	.31	.30	.20	.31	.28	.20	.30	.28
		TOT	125	30.53	43.89	.31	.25	.30	.30	.21	.28	.28	.22	.27	.27	.23	.27
	A-2	MST	125	29.02	20.56	.21	.05	.13	.21	.00	.10	.21	.02	.12	.19	.04	.12
		LST	125	27.65	13.99	.27	.38	.36	.28	.37	.36	.24	.35	.33	.22	.35	.33
		TOT	125	56.66	26.42	.30	.24	.29	.31	.20	.27	.29	.20	.27	.26	.22	.26
	A-3	MST	125	.39	.49	.16	-0.01	.07	.17	-0.05	.05	.17	-0.03	.06	.14	-0.01	.06
		LST	125	.32	.47	.15	.24	.22	.17	.24	.23	.13	.26	.22	.12	.24	.20
		TOT	125	.71	.63	.24	.17	.22	.26	.14	.21	.23	.17	.22	.20	.17	.20
	A-4	MST	125	.29	.64	.27	.12	.20	.27	.08	.18	.26	.11	.19	.25	.11	.19
		LST	125	.27	.54	.19	.29	.26	.20	.28	.27	.17	.30	.26	.16	.28	.25
		TOT	125	.56	.89	.31	.26	.31	.31	.23	.29	.29	.26	.30	.28	.25	.29
24	A-1	MST	126	52.14	39.58	.27	.18	.24	.30	.20	.27	.25	.18	.24	.23	.16	.21
		LST	126	37.35	35.56	.22	.41	.36	.23	.44	.37	.20	.40	.34	.20	.38	.33
		TOT	126	89.49	56.02	.34	.39	.40	.35	.42	.43	.30	.38	.38	.29	.35	.36
	A-2	MST	126	56.11	32.80	.29	.19	.25	.31	.21	.28	.27	.19	.25	.26	.17	.23
		LST	126	42.18	26.50	.10	.32	.24	.10	.35	.26	.08	.32	.23	.08	.29	.22
		TOT	126	98.29	42.04	.29	.35	.35	.30	.39	.38	.26	.35	.34	.25	.32	.31
	A-3	MST	126	.69	.46	.28	.17	.23	.30	.20	.27	.26	.17	.23	.25	.15	.21
		LST	126	.54	.50	.03	.24	.16	.03	.28	.18	.03	.25	.17	.02	.23	.15
		TOT	126	1.23	.66	.22	.30	.29	.24	.35	.32	.20	.31	.29	.19	.28	.26
	A-4	MST	126	.68	.48	.27	.15	.22	.29	.17	.24	.24	.15	.21	.23	.13	.19
		LST	126	.51	.56	.15	.33	.27	.15	.36	.29	.13	.33	.27	.13	.31	.25
		TOT	126	1.19	.74	.29	.35	.35	.30	.38	.38	.26	.35	.34	.25	.32	.31

FORM A (Continued)

25	A-1	MST	126	34.41	19.55	.20	.17	.20	.19	.20	.21	.18	.20	.21	.18	.17	.19
		LST	126	37.13	28.97	.24	.42	.37	.21	.43	.36	.23	.47	.40	.25	.43	.39
		TOT	126	71.55	39.47	.27	.39	.37	.25	.41	.37	.26	.44	.39	.27	.40	.38
	A-2	MST	126	40.93	14.53	.23	.14	.19	.22	.17	.21	.22	.17	.21	.22	.14	.19
		LST	126	42.33	19.46	.22	.39	.34	.20	.40	.34	.21	.45	.38	.23	.40	.36
		TOT	126	83.25	25.94	.30	.37	.36	.27	.39	.37	.28	.43	.40	.29	.38	.38
	A-3	MST	126	.87	.33	.23	.14	.19	.22	.17	.21	.22	.17	.21	.22	.14	.19
		LST	126	.54	.50	.19	.31	.28	.18	.32	.28	.18	.38	.32	.20	.33	.30
		TOT	126	1.41	.62	.28	.32	.33	.26	.35	.34	.26	.40	.37	.28	.34	.35
	A-4	MST	126	.86	.39	.20	.15	.19	.19	.18	.20	.19	.19	.21	.19	.16	.19
		LST	126	.46	.64	.23	.39	.35	.21	.40	.34	.22	.45	.38	.24	.41	.37
		TOT	126	1.32	.82	.28	.37	.36	.25	.40	.36	.26	.44	.39	.28	.39	.38
26	A-1	MST	125	25.88	32.14	.28	.21	.26	.28	.19	.25	.25	.17	.23	.27	.19	.25
		LST	125	30.59	39.81	.30	.40	.38	.29	.43	.40	.31	.42	.40	.29	.38	.38
		TOT	125	56.47	53.93	.39	.42	.44	.38	.43	.45	.37	.41	.43	.38	.40	.43
	A-2	MST	125	33.98	23.40	.26	.15	.21	.26	.12	.20	.22	.10	.17	.24	.13	.19
		LST	125	41.53	30.12	.32	.41	.40	.31	.45	.42	.32	.43	.42	.31	.39	.39
		TOT	125	75.51	38.43	.40	.41	.44	.40	.42	.45	.38	.40	.43	.39	.38	.42
	A-3	MST	125	.46	.50	.25	.13	.20	.26	.10	.19	.22	.08	.16	.23	.11	.18
		LST	125	.54	.50	.31	.40	.39	.30	.43	.41	.31	.43	.41	.30	.38	.38
		TOT	125	1.01	.70	.40	.37	.42	.40	.38	.43	.38	.36	.40	.38	.35	.40
	A-4	MST	125	.43	.56	.28	.19	.25	.27	.17	.24	.24	.15	.21	.26	.18	.23
		LST	125	.48	.61	.29	.39	.38	.29	.42	.39	.30	.41	.40	.29	.38	.37
		TOT	125	.91	.88	.38	.40	.42	.37	.40	.43	.36	.38	.41	.37	.38	.41
27	A-1	MST	126	6.83	23.59	.27	.23	.27	.25	.25	.27	.29	.28	.31	.30	.24	.29
		LST	126	7.37	19.39	.20	.16	.19	.18	.19	.20	.19	.19	.21	.21	.19	.22
		TOT	126	14.20	33.19	.31	.26	.30	.28	.29	.31	.32	.31	.34	.34	.28	.33
	A-2	MST	126	23.57	10.66	.26	.25	.28	.25	.26	.28	.28	.29	.31	.27	.26	.29
		LST	126	24.70	14.25	.17	.14	.17	.15	.17	.17	.17	.17	.19	.20	.16	.19
		TOT	126	48.27	18.36	.28	.26	.29	.26	.29	.30	.29	.30	.33	.31	.28	.32
	A-3	MST	126	.31	.46	.24	.27	.28	.22	.29	.29	.27	.32	.33	.26	.28	.30
		LST	126	.29	.45	.09	.08	.10	.07	.09	.09	.11	.11	.12	.13	.11	.13
		TOT	126	.60	.58	.27	.28	.30	.23	.30	.30	.30	.34	.35	.30	.31	.34
	A-4	MST	126	.16	.66	.23	.20	.23	.20	.23	.23	.25	.26	.28	.26	.21	.26
		LST	126	.21	.57	.19	.16	.19	.17	.18	.19	.19	.19	.21	.20	.19	.21
		TOT	126	.37	.90	.28	.25	.29	.25	.28	.29	.30	.31	.34	.32	.28	.32

FORM A (Continued)

28	A-1	MST	123	17.24	16.29	.41	.39	.43	.45	.37	.45	.35	.35	.39	.33	.36	.38
		LST	125	5.50	19.91	.25	.28	.29	.27	.26	.29	.20	.23	.23	.20	.23	.24
		TOT	123	22.76	30.32	.38	.39	.42	.42	.37	.43	.32	.33	.36	.31	.35	.36
	A-2	MST	123	32.60	16.60	.35	.30	.35	.41	.29	.37	.29	.27	.31	.27	.28	.30
		LST	125	23.24	10.35	.16	.23	.22	.16	.22	.21	.11	.17	.16	.13	.19	.18
		TOT	123	55.86	21.55	.34	.34	.37	.39	.32	.39	.27	.29	.31	.26	.31	.31
	A-3	MST	123	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		LST	125	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		TOT	123	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
	A-4	MST	123	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		LST	125	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		TOT	123	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
29	A-1	MST	125	12.34	26.38	.36	.28	.35	.36	.27	.34	.36	.26	.34	.37	.26	.34
		LST	125	22.89	31.27	.24	.28	.28	.23	.30	.29	.27	.27	.30	.25	.26	.28
		TOT	125	35.23	44.07	.39	.37	.41	.37	.37	.41	.40	.35	.41	.40	.34	.40
	A-2	MST	125	25.91	17.78	.25	.20	.24	.24	.19	.24	.26	.19	.25	.26	.19	.24
		LST	125	34.14	20.46	.27	.33	.33	.26	.35	.34	.30	.33	.35	.29	.31	.33
		TOT	125	60.06	28.43	.35	.36	.39	.34	.37	.39	.38	.36	.40	.37	.34	.39
	A-3	MST	125	.33	.47	.19	.16	.18	.18	.15	.18	.20	.15	.20	.20	.15	.19
		LST	125	.54	.50	.30	.35	.35	.28	.37	.36	.33	.34	.37	.31	.33	.35
		TOT	125	.86	.70	.34	.35	.37	.32	.37	.38	.37	.35	.40	.36	.33	.38
	A-4	MST	125	.27	.56	.36	.28	.34	.36	.27	.34	.36	.27	.34	.37	.26	.34
		LST	125	.46	.63	.25	.27	.28	.24	.29	.29	.28	.26	.30	.26	.25	.28
		TOT	125	.74	.90	.39	.36	.41	.38	.37	.41	.41	.35	.42	.41	.34	.40
30	A-1	MST	126	47.89	49.70	.36	.21	.29	.40	.16	.30	.29	.14	.23	.27	.18	.24
		LST	126	20.86	35.96	.32	.41	.40	.31	.37	.38	.32	.36	.38	.32	.38	.39
		TOT	126	68.75	69.59	.42	.36	.42	.45	.31	.41	.37	.29	.36	.36	.32	.37
	A-2	MST	126	56.17	41.29	.35	.21	.30	.40	.17	.30	.29	.14	.23	.27	.18	.24
		LST	126	31.03	16.23	.32	.38	.38	.31	.37	.38	.34	.38	.39	.32	.38	.39
		TOT	126	87.21	46.28	.42	.32	.40	.47	.28	.40	.37	.26	.34	.35	.29	.35
	A-3	MST	126	.63	.48	.36	.22	.31	.41	.18	.31	.30	.15	.24	.28	.19	.25
		LST	126	.51	.50	.29	.32	.33	.29	.32	.34	.31	.34	.36	.29	.33	.34
		TOT	126	1.13	.72	.44	.37	.44	.48	.34	.44	.41	.34	.41	.39	.36	.40
	A-4	MST	126	.60	.54	.40	.22	.32	.43	.18	.32	.32	.17	.26	.33	.19	.27
		LST	126	.42	.65	.34	.40	.41	.34	.38	.40	.35	.38	.41	.34	.39	.40
		TOT	126	1.02	.94	.46	.40	.46	.48	.37	.46	.43	.36	.43	.42	.38	.43

FORM A (Continued)

31	A-1	MST	126	42.52	37.07	.40	.23	.33	.39	.23	.33	.37	.20	.31	.38	.21	.31
		LST	126	23.71	26.03	.14	.20	.18	.16	.18	.19	.14	.22	.20	.12	.21	.18
		TOT	126	66.24	50.14	.36	.27	.34	.37	.26	.34	.35	.26	.33	.34	.26	.32
	A-2	MST	126	51.68	28.10	.42	.25	.35	.41	.26	.36	.39	.23	.33	.40	.23	.33
		LST	126	33.78	18.94	.14	.15	.16	.15	.16	.17	.12	.22	.19	.13	.19	.18
		TOT	126	85.46	36.48	.39	.27	.35	.39	.28	.36	.36	.29	.35	.37	.27	.35
	A-3	MST	126	.70	.46	.43	.25	.36	.41	.26	.36	.39	.23	.33	.41	.22	.33
		LST	126	.54	.50	.15	.15	.16	.16	.17	.18	.13	.23	.21	.14	.20	.19
		TOT	126	1.24	.73	.37	.26	.34	.37	.28	.35	.34	.30	.35	.35	.28	.34
	A-4	MST	126	.64	.58	.37	.20	.30	.35	.20	.30	.35	.18	.28	.36	.18	.29
		LST	126	.51	.56	.19	.22	.22	.20	.23	.24	.18	.28	.26	.18	.25	.24
		TOT	126	1.15	.92	.35	.26	.33	.35	.27	.33	.33	.28	.34	.34	.27	.33
32	A-1	MST	126	16.12	21.89	.18	.20	.21	.15	.22	.21	.21	.20	.23	.25	.22	.26
		LST	126	8.91	20.34	.10	.23	.19	.11	.21	.18	.13	.22	.19	.12	.27	.23
		TOT	126	25.03	35.31	.17	.26	.24	.16	.25	.23	.20	.25	.25	.22	.29	.29
	A-2	MST	126	30.45	11.30	.18	.20	.21	.16	.22	.21	.22	.21	.23	.25	.23	.26
		LST	126	23.29	9.03	.09	.18	.15	.10	.17	.15	.13	.20	.18	.13	.25	.21
		TOT	126	53.74	17.40	.16	.22	.21	.15	.23	.21	.21	.24	.25	.23	.27	.28
	A-3	MST	126	.71	.46	.21	.22	.24	.19	.24	.24	.25	.22	.25	.28	.24	.28
		LST	126	.44	.50	.01	.03	.02	.01	.03	.03	.05	.08	.07	.06	.10	.09
		TOT	126	1.14	.78	.13	.15	.15	.12	.16	.15	.17	.18	.19	.20	.20	.22
	A-4	MST	126	.43	.89	.18	.21	.21	.16	.23	.21	.21	.21	.23	.25	.23	.26
		LST	126	.22	.78	.06	.14	.11	.07	.13	.11	.10	.17	.15	.10	.21	.18
		TOT	126	.65	1.43	.14	.20	.19	.14	.21	.19	.19	.22	.23	.21	.25	.26
33	A-1	MST	126	31.97	28.42	.20	.10	.15	.21	.10	.16	.21	.14	.19	.23	.14	.19
		LST	126	18.74	25.25	.07	.28	.21	.11	.23	.19	.08	.26	.20	.05	.32	.22
		TOT	126	50.71	43.78	.17	.23	.22	.20	.20	.21	.18	.24	.24	.17	.28	.25
	A-2	MST	126	41.79	19.80	.18	.10	.15	.21	.10	.16	.20	.14	.18	.20	.14	.18
		LST	126	28.87	12.45	.15	.33	.27	.18	.28	.26	.12	.30	.24	.11	.36	.28
		TOT	126	70.65	26.22	.21	.23	.24	.24	.20	.24	.21	.25	.25	.20	.27	.27
	A-3	MST	126	.83	.38	.18	.07	.13	.18	.08	.13	.20	.12	.17	.22	.11	.17
		LST	126	.67	.47	.18	.30	.27	.23	.25	.26	.17	.27	.25	.15	.33	.28
		TOT	126	1.50	.65	.24	.25	.27	.27	.22	.27	.24	.26	.28	.24	.30	.30
	A-4	MST	126	.71	.67	.19	.08	.14	.18	.09	.14	.21	.13	.18	.24	.12	.19
		LST	126	.54	.72	.13	.32	.26	.17	.27	.25	.13	.29	.24	.10	.36	.27
		TOT	126	1.25	1.10	.21	.26	.25	.22	.23	.25	.21	.27	.27	.22	.31	.29

FORM A (Continued)

34	A-1	MST	126	46.73	33.00	.37	.35	.39	.38	.36	.41	.35	.37	.39	.36	.34	.38
		LST	126	35.46	35.18	.37	.48	.47	.38	.48	.48	.36	.46	.46	.35	.45	.45
		TOT	126	82.19	52.63	.48	.54	.56	.50	.55	.57	.46	.54	.55	.46	.52	.54
	A-2	MST	126	49.59	28.40	.34	.29	.34	.36	.30	.36	.33	.31	.35	.33	.29	.34
		LST	126	40.33	25.43	.31	.39	.38	.31	.40	.39	.28	.37	.37	.28	.36	.36
		TOT	126	89.91	39.19	.45	.46	.50	.46	.48	.52	.42	.47	.49	.42	.44	.48
	A-3	MST	126	.63	.48	.33	.28	.33	.35	.29	.35	.32	.29	.33	.32	.27	.32
		LST	126	.41	.49	.17	.17	.18	.16	.19	.19	.14	.17	.17	.14	.15	.16
		TOT	126	1.05	.66	.36	.33	.37	.38	.35	.40	.34	.34	.37	.34	.31	.35
	A-4	MST	126	.62	.52	.36	.34	.38	.38	.35	.39	.35	.35	.38	.35	.33	.37
		LST	126	.37	.56	.24	.28	.28	.25	.28	.29	.23	.27	.27	.22	.25	.26
		TOT	126	.99	.78	.41	.42	.45	.43	.43	.47	.39	.42	.45	.39	.40	.43
35	A-1	MST	126	36.94	35.70	.49	.38	.47	.49	.36	.46	.48	.35	.45	.48	.35	.45
		LST	126	23.24	24.90	.44	.49	.51	.44	.48	.50	.40	.49	.49	.41	.49	.50
		TOT	126	60.17	48.09	.59	.54	.61	.59	.51	.60	.56	.51	.59	.57	.52	.59
	A-2	MST	126	45.19	25.20	.49	.38	.46	.50	.35	.46	.47	.35	.44	.46	.35	.44
		LST	126	31.40	13.92	.39	.46	.47	.39	.46	.47	.36	.45	.45	.37	.46	.46
		TOT	126	76.60	31.12	.57	.51	.58	.58	.49	.58	.54	.48	.56	.54	.49	.56
	A-3	MST	126	.63	.48	.42	.33	.40	.43	.30	.39	.41	.30	.39	.40	.30	.37
		LST	126	.76	.43	.37	.47	.46	.39	.47	.47	.33	.48	.45	.33	.47	.45
		TOT	126	1.40	.66	.55	.55	.59	.57	.52	.60	.52	.53	.58	.51	.53	.57
	A-4	MST	126	.53	.67	.57	.44	.54	.57	.42	.53	.54	.40	.51	.55	.41	.51
		LST	126	.71	.55	.43	.50	.51	.44	.48	.51	.39	.50	.50	.40	.50	.50
		TOT	126	1.25	.97	.64	.58	.66	.65	.56	.66	.60	.56	.63	.61	.57	.64
36	A-1	MST	124	21.68	29.49	.38	.28	.35	.37	.27	.35	.38	.24	.34	.37	.26	.34
		LST	125	12.62	23.00	.31	.43	.41	.32	.39	.39	.28	.40	.38	.29	.44	.41
		TOT	124	34.48	41.99	.43	.43	.46	.43	.40	.46	.42	.39	.45	.42	.42	.46
	A-2	MST	124	32.62	17.85	.37	.29	.35	.37	.28	.35	.38	.26	.34	.36	.28	.34
		LST	125	27.15	12.03	.28	.41	.38	.29	.38	.37	.26	.38	.36	.26	.41	.37
		TOT	124	59.90	23.53	.42	.43	.46	.42	.40	.45	.41	.39	.44	.41	.42	.45
	A-3	MST	124	.50	.50	.36	.30	.35	.36	.30	.35	.37	.27	.35	.36	.29	.35
		LST	125	.67	.47	.22	.32	.30	.22	.30	.29	.19	.31	.28	.19	.31	.29
		TOT	124	1.18	.73	.38	.41	.43	.38	.39	.42	.38	.38	.42	.37	.40	.42
	A-4	MST	124	.32	.76	.37	.27	.34	.37	.26	.34	.38	.24	.33	.37	.26	.34
		LST	125	.58	.66	.30	.41	.39	.30	.38	.38	.27	.39	.37	.28	.41	.39
		TOT	124	.90	1.15	.41	.41	.44	.41	.38	.43	.40	.38	.43	.40	.41	.44

FORM A (Continued)

37	A-1	MST	126	36.91	29.51	.26	.14	.21	.25	.14	.21	.20	.17	.20	.23	.14	.19
		LST	125	28.34	35.47	.30	.36	.36	.30	.35	.36	.27	.37	.36	.29	.35	.36
		TOT	125	65.09	52.09	.35	.32	.36	.35	.32	.36	.30	.35	.36	.33	.32	.35
	A-2	MST	126	44.25	22.50	.24	.14	.20	.25	.14	.21	.20	.18	.20	.20	.14	.19
		LST	125	37.13	23.83	.24	.31	.30	.24	.30	.29	.22	.32	.30	.23	.30	.29
		TOT	125	81.25	35.63	.31	.30	.33	.32	.29	.33	.27	.32	.33	.28	.29	.31
	A-3	MST	126	.63	.48	.18	.09	.14	.20	.08	.15	.15	.14	.16	.15	.10	.13
		LST	125	.53	.50	.21	.27	.26	.21	.25	.26	.19	.28	.27	.19	.26	.25
		TOT	125	1.15	.75	.25	.23	.26	.27	.22	.27	.22	.27	.28	.22	.24	.25
	A-4	MST	126	.60	.55	.21	.09	.16	.22	.09	.16	.17	.14	.17	.18	.10	.15
		LST	125	.46	.63	.28	.34	.34	.29	.32	.34	.25	.36	.34	.26	.34	.33
		TOT	125	1.05	.95	.30	.27	.31	.31	.26	.31	.26	.31	.32	.28	.28	.31
38	A-1	MST	126	21.49	15.18	.34	.22	.29	.34	.22	.30	.39	.21	.32	.37	.22	.31
		LST	125	23.38	36.25	.26	.35	.34	.27	.35	.34	.32	.33	.36	.27	.38	.36
		TOT	125	45.14	42.00	.34	.38	.39	.35	.38	.40	.42	.36	.42	.36	.40	.42
	A-2	MST	126	30.19	10.06	.27	.13	.21	.28	.13	.22	.35	.12	.25	.31	.13	.23
		LST	125	34.71	24.30	.26	.35	.34	.27	.34	.34	.32	.33	.36	.27	.37	.36
		TOT	125	65.08	27.38	.33	.35	.37	.35	.35	.38	.41	.34	.41	.35	.38	.40
	A-3	MST	126	.74	.44	.20	.10	.16	.20	.11	.17	.28	.10	.20	.24	.10	.18
		LST	125	.49	.50	.27	.31	.31	.27	.31	.32	.32	.30	.34	.28	.33	.34
		TOT	125	1.23	.69	.32	.28	.32	.32	.29	.34	.41	.28	.37	.35	.30	.35
	A-4	MST	126	.72	.48	.26	.15	.21	.26	.15	.22	.33	.14	.25	.30	.14	.23
		LST	125	.28	.79	.25	.34	.33	.25	.34	.33	.32	.31	.35	.27	.37	.36
		TOT	125	1.01	1.00	.32	.34	.36	.32	.34	.36	.41	.31	.39	.36	.36	.39
39	A-1	MST	126	13.76	26.67	.18	.19	.20	.20	.19	.21	.17	.19	.19	.16	.19	.19
		LST	126	18.93	26.94	.31	.37	.37	.31	.33	.35	.31	.35	.36	.30	.37	.37
		TOT	126	32.69	44.94	.29	.33	.34	.30	.31	.33	.28	.32	.33	.28	.33	.34
	A-2	MST	126	26.40	17.20	.11	.15	.14	.13	.15	.15	.11	.14	.14	.10	.14	.14
		LST	126	29.34	13.16	.35	.41	.41	.34	.38	.39	.34	.40	.41	.33	.41	.41
		TOT	126	55.75	24.68	.26	.32	.32	.27	.30	.32	.25	.31	.31	.25	.32	.31
	A-3	MST	126	.28	.45	-0.04	.04	.00	-0.02	.05	.02	-0.03	.04	.01	-0.04	.04	.01
		LST	126	.71	.45	.34	.42	.41	.33	.39	.39	.32	.41	.41	.32	.41	.41
		TOT	126	.99	.68	.20	.30	.28	.20	.29	.27	.19	.30	.27	.18	.30	.27
	A-4	MST	126	.13	.64	.16	.18	.18	.17	.18	.19	.14	.17	.17	.14	.17	.17
		LST	126	.63	.62	.32	.37	.38	.32	.34	.36	.33	.35	.37	.31	.38	.38
		TOT	126	.76	1.06	.29	.33	.34	.30	.31	.33	.28	.31	.32	.27	.33	.33

FORM A (Continued)

40	A-1	MST	125	14.52	25.37	.18	.02	.10	.13	.03	.08	.20	.02	.11	.23	.01	.12
		LST	126	14.22	23.56	.17	.18	.19	.17	.18	.19	.19	.22	.23	.19	.21	.22
		TOT	125	28.67	37.69	.24	.13	.19	.20	.14	.18	.27	.16	.23	.29	.15	.23
	A-2	MST	125	27.92	15.86	.18	.02	.10	.15	.04	.09	.18	.03	.11	.21	.02	.11
		LST	126	26.37	11.96	.15	.17	.18	.16	.18	.18	.15	.22	.21	.16	.20	.20
		TOT	125	54.22	21.55	.23	.12	.18	.20	.13	.18	.23	.15	.21	.25	.14	.21
	A-3	MST	125	.66	.48	.14	-0.01	.06	.08	.01	.05	.18	-0.01	.09	.21	-0.01	.09
		LST	126	.58	.49	.10	.12	.12	.11	.12	.12	.09	.17	.15	.10	.14	.14
		TOT	125	1.23	.71	.18	.09	.14	.15	.10	.13	.20	.13	.18	.23	.11	.17
	A-4	MST	125	.45	.81	.14	-0.00	.07	.09	.02	.05	.18	-0.00	.09	.22	-0.01	.10
		LST	126	.42	.75	.16	.18	.18	.15	.18	.18	.17	.23	.22	.18	.21	.22
		TOT	125	.86	1.19	.21	.12	.18	.17	.14	.17	.25	.15	.21	.28	.14	.22
41	A-1	MST	126	24.90	22.50	.34	.28	.33	.33	.24	.31	.33	.25	.32	.34	.29	.35
		LST	126	22.84	36.75	.38	.52	.50	.39	.50	.50	.38	.51	.49	.35	.51	.48
		TOT	126	47.74	48.57	.45	.52	.53	.45	.50	.52	.44	.50	.52	.43	.52	.53
	A-2	MST	126	35.29	17.76	.37	.31	.36	.36	.27	.34	.36	.28	.35	.38	.33	.38
		LST	126	35.32	25.13	.38	.47	.47	.38	.47	.47	.36	.48	.47	.35	.46	.45
		TOT	126	70.61	33.98	.47	.51	.53	.47	.49	.53	.46	.50	.53	.45	.51	.54
	A-3	MST	126	.50	.50	.24	.22	.25	.25	.18	.23	.25	.19	.24	.25	.24	.27
		LST	126	.48	.50	.35	.44	.44	.35	.45	.44	.34	.45	.44	.32	.43	.42
		TOT	126	.98	.75	.40	.45	.46	.40	.42	.45	.40	.43	.46	.39	.45	.47
	A-4	MST	126	.48	.53	.27	.24	.27	.27	.20	.25	.28	.20	.26	.28	.25	.29
		LST	126	.39	.65	.36	.53	.49	.37	.51	.49	.35	.52	.49	.32	.53	.48
		TOT	126	.87	.93	.40	.51	.50	.41	.47	.49	.40	.48	.49	.38	.51	.50
42	A-1	MST	125	18.58	27.73	.35	.28	.33	.32	.26	.32	.37	.26	.34	.38	.28	.35
		LST	125	14.93	31.64	.35	.40	.41	.32	.39	.39	.35	.40	.41	.37	.42	.43
		TOT	125	33.50	50.72	.41	.40	.44	.37	.39	.41	.42	.39	.44	.44	.41	.46
	A-2	MST	125	29.09	15.78	.34	.28	.33	.31	.27	.32	.37	.25	.34	.37	.28	.35
		LST	125	27.98	15.45	.33	.40	.40	.30	.39	.38	.32	.40	.40	.35	.42	.42
		TOT	125	57.07	26.27	.40	.41	.44	.36	.39	.41	.41	.39	.44	.43	.41	.46
	A-3	MST	125	.63	.48	.33	.27	.32	.30	.26	.31	.37	.24	.32	.36	.27	.34
		LST	125	.63	.48	.29	.36	.36	.26	.35	.34	.29	.36	.36	.31	.37	.38
		TOT	125	1.26	.78	.38	.39	.42	.35	.38	.40	.40	.37	.42	.42	.39	.44
	A-4	MST	125	.49	.73	.35	.28	.33	.31	.27	.32	.37	.26	.34	.38	.28	.35
		LST	125	.38	.86	.35	.39	.40	.32	.37	.38	.34	.39	.40	.37	.40	.42
		TOT	125	.87	1.36	.40	.39	.43	.37	.38	.41	.41	.38	.43	.43	.41	.46

FORM A (Continued)

43	A-1	MST	126	30.13	33.19	.45	.29	.39	.44	.30	.39	.43	.29	.39	.47	.30	.41
		LST	125	8.68	26.96	.30	.39	.38	.27	.39	.36	.24	.42	.38	.28	.40	.38
		TOT	125	38.97	51.12	.45	.40	.46	.43	.40	.45	.41	.41	.45	.46	.40	.47
	A-2	MST	126	41.14	23.64	.44	.30	.39	.43	.30	.39	.43	.30	.39	.46	.30	.41
		LST	125	23.54	10.17	.31	.41	.40	.28	.42	.39	.27	.42	.39	.30	.41	.40
		TOT	125	64.81	29.50	.46	.38	.45	.45	.38	.45	.44	.38	.45	.47	.39	.47
	A-3	MST	126	.54	.50	.35	.18	.28	.35	.18	.28	.36	.18	.29	.38	.19	.29
		LST	125	.73	.44	.34	.47	.45	.33	.46	.44	.28	.48	.43	.32	.46	.44
		TOT	125	1.27	.74	.45	.40	.46	.44	.40	.45	.42	.41	.45	.45	.40	.46
	A-4	MST	126	.35	.78	.46	.32	.41	.45	.33	.42	.44	.32	.41	.48	.33	.43
		LST	125	.62	.68	.31	.41	.40	.29	.41	.38	.26	.44	.39	.30	.42	.40
		TOT	125	.97	1.25	.46	.43	.48	.44	.43	.47	.42	.44	.48	.47	.43	.49
44	A-1	MST	126	11.67	24.44	.18	.09	.14	.16	.09	.13	.21	.10	.17	.23	.09	.17
		LST	126	28.89	36.77	.22	.38	.33	.23	.41	.35	.26	.37	.35	.24	.37	.34
		TOT	126	40.56	48.12	.26	.33	.33	.26	.36	.34	.30	.33	.35	.30	.33	.34
	A-2	MST	126	23.25	16.67	.13	.04	.09	.13	.02	.08	.16	.05	.11	.17	.05	.11
		LST	126	41.77	26.28	.22	.36	.33	.21	.40	.34	.25	.36	.34	.24	.35	.33
		TOT	126	65.02	32.68	.24	.31	.30	.24	.33	.32	.29	.31	.33	.28	.31	.33
	A-3	MST	126	.26	.44	.12	.03	.08	.13	.01	.07	.15	.04	.10	.16	.05	.11
		LST	126	.60	.49	.22	.35	.32	.21	.39	.33	.25	.35	.33	.24	.34	.32
		TOT	126	.87	.68	.23	.27	.28	.23	.29	.29	.28	.28	.30	.27	.28	.30
	A-4	MST	126	.21	.51	.18	.09	.14	.17	.09	.14	.22	.10	.17	.23	.10	.17
		LST	126	.50	.68	.22	.37	.33	.23	.40	.35	.25	.36	.35	.23	.36	.33
		TOT	126	.71	.92	.26	.32	.32	.26	.34	.33	.31	.32	.35	.30	.32	.34
45	A-1	MST	126	1.52	21.52	.11	.07	.10	.10	.12	.12	.17	.15	.17	.20	.11	.17
		LST	125	6.71	20.93	.06	.18	.14	.09	.16	.14	.10	.20	.17	.08	.22	.18
		TOT	125	8.11	34.55	.11	.16	.15	.12	.17	.16	.16	.21	.21	.18	.21	.21
	A-2	MST	126	20.66	9.75	.07	.03	.06	.06	.08	.08	.13	.12	.14	.16	.08	.13
		LST	125	23.86	11.39	.08	.25	.19	.11	.22	.19	.10	.26	.21	.09	.28	.22
		TOT	125	44.50	17.44	.09	.18	.15	.11	.19	.17	.14	.24	.21	.15	.23	.21
	A-3	MST	126	.64	.48	.12	.07	.10	.10	.12	.12	.17	.15	.17	.21	.11	.17
		LST	125	.54	.50	.07	.18	.14	.09	.16	.14	.10	.20	.17	.09	.23	.18
		TOT	125	1.18	.79	.11	.16	.15	.12	.17	.16	.17	.22	.21	.18	.21	.22
	A-4	MST	126	.29	.96	.12	.07	.10	.10	.12	.12	.17	.15	.17	.21	.11	.17
		LST	125	.09	1.00	.07	.18	.14	.09	.16	.14	.10	.20	.17	.09	.23	.18
		TOT	125	.37	1.59	.11	.16	.15	.12	.17	.16	.17	.22	.21	.18	.21	.22

FORM A (Continued)

46	A-1	MST	126	18.35	24.82	.17	.15	.17	.13	.17	.17	.21	.16	.20	.19	.15	.18
		LST	126	16.45	31.43	.26	.36	.34	.26	.37	.35	.31	.37	.38	.29	.37	.37
		TOT	126	34.80	44.74	.27	.34	.34	.26	.36	.34	.33	.35	.38	.31	.34	.36
	A-2	MST	126	27.70	13.41	.15	.17	.18	.13	.20	.18	.19	.19	.21	.18	.17	.19
		LST	126	33.62	24.10	.24	.34	.32	.24	.36	.33	.31	.36	.37	.28	.35	.35
		TOT	126	61.32	29.73	.26	.35	.34	.25	.38	.35	.33	.38	.40	.31	.36	.37
	A-3	MST	126	.67	.47	.16	.18	.19	.13	.21	.19	.20	.20	.22	.19	.17	.20
		LST	126	.50	.50	.20	.29	.27	.19	.32	.28	.27	.33	.33	.25	.31	.31
		TOT	126	1.17	.76	.23	.30	.29	.21	.34	.30	.30	.34	.36	.28	.31	.33
	A-4	MST	126	.62	.59	.16	.16	.17	.13	.18	.17	.20	.16	.20	.19	.15	.18
		LST	126	.25	.83	.25	.35	.33	.25	.36	.34	.30	.36	.37	.28	.36	.36
		TOT	126	.87	1.15	.26	.33	.33	.25	.35	.33	.32	.35	.37	.30	.34	.35
47	A-1	MST	126	26.90	28.80	.27	.16	.23	.26	.16	.22	.29	.12	.22	.27	.13	.21
		LST	125	37.43	39.12	.36	.43	.43	.34	.45	.43	.34	.41	.42	.35	.41	.42
		TOT	125	64.08	54.20	.40	.40	.43	.38	.41	.43	.41	.36	.42	.40	.37	.42
	A-2	MST	126	35.62	20.76	.26	.13	.21	.28	.13	.22	.30	.09	.20	.25	.11	.19
		LST	125	44.78	25.96	.36	.41	.42	.34	.44	.43	.35	.41	.42	.36	.39	.41
		TOT	125	80.21	35.63	.41	.38	.43	.41	.40	.44	.43	.35	.43	.41	.35	.41
	A-3	MST	126	.38	.49	.15	.06	.11	.16	.04	.11	.22	.02	.12	.18	.06	.12
		LST	125	.64	.48	.35	.40	.41	.32	.44	.42	.33	.40	.41	.35	.38	.40
		TOT	125	1.02	.72	.34	.31	.35	.32	.32	.35	.38	.29	.36	.36	.30	.36
	A-4	MST	126	.36	.53	.20	.11	.16	.19	.10	.16	.25	.07	.17	.23	.10	.17
		LST	125	.57	.62	.33	.39	.39	.31	.41	.40	.32	.38	.39	.32	.37	.39
		TOT	125	.92	.90	.35	.34	.37	.33	.35	.37	.37	.31	.37	.36	.32	.37
48	A-1	MST	125	5.14	14.52	.21	.04	.12	.17	.07	.13	.18	.03	.11	.20	.00	.09
		LST	125	10.25	15.13	.31	.25	.30	.28	.26	.30	.29	.23	.28	.29	.20	.26
		TOT	125	15.38	24.53	.32	.18	.26	.28	.20	.26	.28	.16	.23	.30	.12	.22
	A-2	MST	125	22.24	6.50	.15	.10	.13	.13	.11	.13	.10	.08	.10	.14	.08	.11
		LST	125	29.26	13.04	.43	.33	.41	.38	.36	.40	.41	.32	.40	.43	.29	.38
		TOT	125	51.50	14.82	.45	.34	.42	.39	.37	.41	.41	.31	.39	.44	.29	.39
	A-3	MST	125	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		LST	125	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		TOT	125	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
	A-4	MST	125	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		LST	125	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		TOT	125	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00

FORM B

Itm	T-y-p-e	N	Mean	A.D.	ALGORITHM-1 MST	ALGORITHM-1 LST	ALGORITHM-1 TOT	ALGORITHM-2 MST	ALGORITHM-2 LST	ALGORITHM-2 TOT	ALGORITHM-3 MST	ALGORITHM-3 LST	ALGORITHM-3 TOT	ALGORITM-4 MST	ALGORITM-4 LST	ALGORITM-4 TOT
1	A-1 MST	123	22.38	20.72	-0.07	-0.14	-0.12	-0.07	-0.16	-0.14	-0.03	-0.14	-0.11	-0.05	-0.14	-0.11
	LST	123	36.96	39.49	.12	.23	.20	.10	.23	.19	.09	.21	.18	.10	.24	.20
	TOT	123	59.34	47.86	.07	.13	.12	.05	.12	.10	.06	.11	.11	.06	.14	.12
	A-2 MST	123	29.91	15.49	-0.08	-0.16	-0.14	-0.08	-0.18	-0.15	-0.04	-0.15	-0.12	-0.06	-0.15	-0.12
	LST	123	45.43	27.68	.14	.24	.21	.11	.24	.20	.09	.22	.19	.11	.24	.20
	TOT	123	75.34	33.50	.08	.12	.11	.05	.11	.10	.06	.11	.10	.06	.13	.11
	A-3 MST	123	.66	.47	-0.07	-0.16	-0.14	-0.07	-0.18	-0.15	-0.03	-0.15	-0.11	-0.05	-0.15	-0.12
	LST	123	.64	.48	.14	.24	.21	.11	.24	.20	.09	.22	.19	.11	.24	.20
	TOT	123	1.30	.71	.05	.05	.05	.03	.04	.04	.04	.05	.05	.04	.06	.06
	A-4 MST	123	.65	.49	-0.06	-0.15	-0.12	-0.06	-0.17	-0.14	-0.02	-0.14	-0.10	-0.04	-0.14	-0.10
	LST	123	.50	.74	.12	.23	.20	.09	.22	.18	.08	.21	.18	.09	.24	.19
	TOT	123	1.15	.96	.06	.10	.09	.04	.09	.07	.05	.09	.08	.05	.11	.09
2	A-1 MST	122	28.26	36.35	.11	.18	.16	.11	.16	.15	.07	.19	.16	.12	.20	.18
	LST	121	12.21	38.99	.04	.17	.13	.06	.14	.12	.05	.18	.15	.01	.19	.12
	TOT	121	40.24	67.11	.08	.20	.16	.10	.17	.15	.07	.21	.17	.07	.22	.17
	A-2 MST	122	43.91	28.39	.10	.16	.14	.11	.14	.14	.09	.17	.15	.11	.18	.17
	LST	121	30.90	20.26	.09	.23	.19	.11	.22	.19	.12	.27	.24	.07	.27	.20
	TOT	121	74.61	40.68	.11	.23	.19	.13	.21	.19	.12	.25	.23	.11	.27	.22
	A-3 MST	122	.57	.49	.10	.16	.15	.11	.14	.14	.10	.16	.15	.11	.19	.17
	LST	121	.45	.50	.10	.24	.19	.12	.23	.20	.13	.28	.25	.08	.27	.21
	TOT	121	1.02	.82	.12	.24	.21	.14	.22	.20	.14	.27	.25	.12	.28	.23
	A-4 MST	122	.47	.68	.10	.17	.16	.10	.16	.15	.07	.19	.16	.11	.20	.18
	LST	121	.27	.75	.06	.18	.14	.08	.16	.14	.08	.20	.17	.03	.20	.14
	TOT	121	.74	1.26	.09	.20	.17	.11	.18	.16	.09	.22	.19	.08	.23	.18
3	A-1 MST	124	28.94	26.01	.36	.25	.32	.32	.25	.31	.30	.23	.30	.36	.24	.32
	LST	124	12.85	25.65	.00	.09	.06	.00	.10	.06	-0.10	.17	.06	-0.03	.13	.07
	TOT	124	41.79	39.39	.24	.23	.25	.21	.23	.25	.13	.26	.24	.22	.24	.26
	A-2 MST	124	36.79	14.77	.37	.25	.33	.33	.25	.32	.32	.23	.31	.37	.24	.32
	LST	124	26.74	16.14	-0.03	.06	.02	-0.03	.06	.03	-0.13	.14	.02	-0.06	.10	.03
	TOT	124	63.53	23.10	.22	.20	.22	.20	.21	.22	.11	.24	.22	.19	.22	.23
	A-3 MST	124	.71	.45	.36	.25	.32	.33	.25	.31	.32	.21	.30	.36	.23	.31
	LST	124	.35	.48	-0.00	.03	.02	-0.01	.04	.02	-0.12	.13	.03	-0.05	.09	.03
	TOT	124	1.06	.69	.23	.19	.22	.21	.19	.22	.12	.23	.22	.20	.21	.23
	A-4 MST	124	.68	.53	.37	.27	.34	.33	.27	.32	.30	.24	.31	.37	.25	.33
	LST	124	.32	.53	.03	.09	.07	.02	.09	.07	-0.08	.17	.08	-0.01	.13	.08
	TOT	124	1.00	.82	.26	.23	.26	.23	.23	.25	.15	.27	.25	.23	.24	.26

FORM B (Continued)

4	A-1	MST	124	22.84	26.87	.12	.03	.07	.15	.01	.08	.15	-0.03	.05	.09	-0.01	.04
		LST	124	43.99	36.74	.10	.34	.26	.11	.35	.27	.05	.31	.23	.06	.31	.22
		TOT	124	66.83	42.58	.16	.31	.27	.19	.31	.28	.14	.25	.23	.11	.26	.22
	A-2	MST	124	29.85	22.37	.11	.02	.06	.14	.01	.07	.14	-0.03	.04	.08	-0.01	.03
		LST	124	49.78	26.99	.10	.30	.24	.11	.31	.25	.05	.28	.21	.06	.28	.21
		TOT	124	79.64	32.04	.16	.27	.24	.20	.26	.26	.14	.21	.21	.11	.23	.19
	A-3	MST	124	.35	.48	.11	.02	.06	.15	.01	.08	.14	-0.03	.05	.09	-0.01	.04
		LST	124	.69	.46	.09	.29	.22	.11	.29	.23	.05	.27	.20	.05	.26	.19
		TOT	124	1.04	.60	.16	.24	.22	.20	.23	.24	.15	.18	.19	.11	.19	.17
	A-4	MST	124	.35	.48	.11	.02	.06	.15	.01	.08	.14	-0.03	.05	.09	-0.01	.04
		LST	124	.65	.54	.10	.35	.26	.11	.35	.27	.05	.31	.23	.06	.31	.23
		TOT	124	1.01	.68	.16	.29	.25	.19	.29	.27	.14	.22	.22	.11	.24	.20
5	A-1	MST	124	23.76	28.83	.07	-0.09	-0.02	.06	-0.11	-0.04	.08	-0.02	.03	.11	-0.05	.03
		LST	124	11.02	18.86	.06	.17	.13	.07	.15	.13	.08	.20	.17	.08	.19	.16
		TOT	124	34.78	38.97	.08	.01	.05	.08	-0.01	.03	.10	.08	.10	.12	.06	.09
	A-2	MST	124	33.27	21.10	.07	-0.11	-0.03	.08	-0.13	-0.05	.11	-0.05	.02	.12	-0.06	.02
		LST	124	27.52	7.63	.12	.25	.21	.11	.27	.22	.09	.26	.22	.13	.24	.21
		TOT	124	60.79	23.25	.10	-0.02	.04	.11	-0.03	.03	.13	.04	.09	.15	.02	.09
	A-3	MST	124	.46	.50	.05	-0.13	-0.06	.05	-0.15	-0.07	.10	-0.06	.01	.10	-0.08	-0.00
		LST	124	.44	.50	.07	.16	.13	.08	.14	.12	.08	.20	.17	.08	.19	.16
		TOT	124	.90	.80	.07	.02	.04	.08	-0.01	.03	.11	.09	.11	.11	.07	.10
	A-4	MST	124	.36	.65	.08	-0.08	-0.01	.07	-0.10	-0.03	.08	-0.01	.03	.12	-0.04	.03
		LST	124	.41	.54	.05	.15	.12	.07	.13	.12	.08	.19	.17	.07	.18	.15
		TOT	124	.77	.97	.08	.03	.06	.09	.01	.04	.10	.10	.11	.12	.07	.10
6	A-1	MST	124	42.71	44.06	.21	.14	.18	.20	.13	.17	.16	.08	.13	.16	.09	.14
		LST	124	37.37	39.74	.36	.37	.40	.36	.39	.41	.31	.34	.38	.33	.33	.36
		TOT	124	80.08	65.89	.36	.31	.36	.35	.32	.36	.30	.26	.32	.31	.26	.31
	A-2	MST	124	51.11	32.83	.18	.10	.15	.19	.09	.14	.15	.06	.11	.13	.06	.10
		LST	124	47.83	31.72	.30	.34	.35	.29	.36	.37	.25	.31	.33	.26	.30	.32
		TOT	124	98.94	47.67	.32	.30	.33	.32	.30	.34	.27	.25	.30	.27	.25	.28
	A-3	MST	124	.65	.48	.19	.11	.15	.19	.10	.15	.15	.07	.11	.13	.07	.11
		LST	124	.61	.49	.29	.33	.34	.28	.35	.35	.24	.30	.32	.25	.29	.31
		TOT	124	1.27	.71	.32	.30	.33	.32	.31	.34	.26	.25	.30	.26	.25	.28
	A-4	MST	124	.59	.61	.21	.14	.19	.20	.13	.18	.16	.09	.14	.16	.10	.14
		LST	124	.56	.60	.36	.37	.39	.35	.39	.41	.30	.34	.37	.32	.33	.36
		TOT	124	1.15	.96	.36	.32	.37	.35	.33	.37	.29	.27	.32	.31	.27	.32

FORM B (Continued)

7	A-1	MST	124	32.51	26.79	.26	.08	.17	.28	.10	.19	.35	.12	.25	.29	.10	.20
		LST	124	41.85	29.13	.42	.48	.49	.39	.49	.49	.37	.49	.50	.40	.49	.50
		TOT	124	74.36	46.77	.41	.35	.41	.40	.36	.42	.43	.37	.46	.41	.37	.43
	A-2	MST	124	37.06	16.84	.23	.08	.16	.26	.09	.18	.33	.11	.24	.26	.10	.19
		LST	124	46.40	22.34	.40	.42	.44	.36	.42	.44	.34	.44	.45	.39	.44	.46
		TOT	124	83.46	31.64	.40	.34	.40	.39	.35	.40	.42	.37	.45	.41	.36	.42
	A-3	MST	124	.79	.41	.20	.06	.13	.22	.06	.14	.30	.08	.20	.23	.07	.15
		LST	124	.59	.49	.31	.28	.31	.27	.28	.31	.26	.31	.34	.31	.31	.34
		TOT	124	1.38	.73	.32	.22	.28	.31	.23	.29	.35	.26	.34	.34	.25	.32
	A-4	MST	124	.75	.52	.23	.05	.14	.24	.07	.16	.32	.09	.22	.26	.07	.17
		LST	124	.53	.60	.36	.39	.41	.34	.40	.41	.33	.41	.43	.35	.42	.43
		TOT	124	1.28	.93	.36	.28	.34	.35	.30	.35	.39	.32	.40	.37	.31	.37
8	A-1	MST	123	11.89	19.77	.31	.30	.33	.29	.29	.32	.29	.26	.32	.27	.28	.31
		LST	122	22.05	25.21	.39	.55	.52	.37	.57	.53	.30	.49	.47	.36	.50	.49
		TOT	122	33.87	39.17	.40	.50	.50	.39	.51	.50	.34	.45	.46	.37	.47	.47
	A-2	MST	123	30.27	11.61	.38	.31	.37	.40	.31	.38	.41	.29	.40	.39	.32	.38
		LST	122	39.67	23.01	.35	.45	.44	.34	.47	.45	.27	.39	.39	.32	.40	.40
		TOT	122	69.86	29.00	.43	.48	.50	.43	.50	.51	.38	.43	.47	.41	.44	.47
	A-3	MST	123	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		LST	122	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		TOT	122	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
	A-4	MST	123	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		LST	122	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		TOT	122	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
9	A-1	MST	124	37.55	35.40	.27	.23	.27	.30	.24	.29	.33	.21	.30	.26	.21	.26
		LST	123	32.66	33.59	.24	.39	.36	.26	.40	.38	.22	.37	.35	.19	.37	.32
		TOT	123	70.47	55.75	.31	.38	.38	.34	.39	.41	.34	.35	.40	.27	.35	.35
	A-2	MST	124	44.98	25.70	.27	.24	.27	.31	.24	.30	.33	.21	.30	.26	.22	.26
		LST	123	41.63	26.59	.23	.42	.37	.25	.43	.39	.22	.40	.37	.18	.39	.34
		TOT	123	86.88	40.74	.31	.42	.41	.35	.43	.44	.34	.38	.42	.27	.38	.37
	A-3	MST	124	.65	.48	.27	.24	.27	.31	.24	.30	.32	.21	.30	.26	.22	.26
		LST	123	.56	.50	.22	.41	.36	.24	.42	.38	.21	.38	.36	.17	.38	.33
		TOT	123	1.21	.75	.30	.42	.40	.35	.42	.43	.33	.38	.42	.27	.38	.37
	A-4	MST	124	.60	.58	.26	.23	.27	.30	.24	.29	.32	.21	.30	.26	.21	.25
		LST	123	.52	.58	.23	.39	.35	.26	.39	.37	.22	.35	.35	.18	.36	.31
		TOT	123	1.12	.92	.30	.38	.38	.34	.38	.40	.33	.35	.40	.26	.35	.35

FORM B (Continued)

10	A-1	MST	122	1.70	10.55	.05	-0.07	-0.02	.07	-0.07	-0.01	.11	-0.11	-0.01	.06	-0.10	-0.03
		LST	122	4.15	11.83	.07	.12	.10	.07	.12	.10	.08	.05	.07	.08	.08	.09
		TOT	122	5.84	18.36	.07	.04	.06	.08	.03	.06	.11	-0.03	.04	.08	-0.00	.04
	A-2	MST	122	23.79	7.67	.19	.09	.15	.25	.10	.18	.21	.11	.18	.21	.08	.15
		LST	122	24.61	13.32	.01	.06	.04	-0.00	.07	.04	.01	.00	.00	.02	.04	.03
		TOT	122	48.39	14.20	.11	.11	.12	.13	.11	.13	.12	.06	.10	.13	.08	.11
	A-3	MST	122	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		LST	122	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		TOT	122	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
	A-4	MST	122	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		LST	122	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		TOT	122	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
11	A-1	MST	123	17.37	20.80	.32	.16	.25	.30	.18	.25	.32	.21	.30	.38	.18	.29
		LST	124	11.85	19.26	.16	.20	.20	.13	.20	.19	.13	.20	.20	.19	.21	.23
		TOT	123	29.17	29.40	.35	.25	.32	.31	.26	.31	.32	.28	.35	.40	.27	.36
	A-2	MST	123	31.90	12.11	.38	.20	.30	.36	.22	.30	.37	.24	.34	.44	.21	.34
		LST	124	24.75	8.39	.15	.20	.19	.12	.19	.18	.11	.19	.18	.18	.20	.21
		TOT	123	56.66	15.14	.39	.27	.35	.35	.28	.34	.36	.30	.38	.45	.28	.39
	A-3	MST	123	.27	.44	.07	.02	.04	.04	.03	.04	.09	.08	.10	.11	.06	.09
		LST	124	.83	.38	.19	.22	.22	.16	.21	.21	.16	.22	.22	.21	.23	.24
		TOT	123	1.10	.55	.19	.17	.20	.16	.18	.19	.20	.22	.24	.24	.21	.25
	A-4	MST	123	.10	.66	.30	.15	.23	.27	.17	.23	.30	.21	.28	.36	.18	.28
		LST	124	.76	.57	.15	.18	.18	.12	.17	.17	.12	.18	.18	.18	.20	.21
		TOT	123	.85	.90	.33	.23	.29	.29	.24	.28	.31	.27	.33	.38	.26	.35
12	A-1	MST	124	54.28	45.59	.39	.22	.32	.42	.21	.33	.37	.21	.32	.36	.23	.32
		LST	124	32.44	22.04	.33	.44	.43	.29	.43	.41	.24	.41	.39	.35	.44	.44
		TOT	124	86.73	54.33	.46	.36	.44	.47	.35	.44	.41	.34	.43	.44	.37	.45
	A-2	MST	124	59.94	36.10	.37	.20	.30	.40	.19	.31	.37	.18	.30	.34	.21	.29
		LST	124	37.54	16.91	.24	.39	.36	.20	.40	.34	.16	.37	.33	.26	.40	.38
		TOT	124	97.48	41.32	.42	.34	.41	.43	.33	.41	.39	.31	.40	.41	.35	.41
	A-3	MST	124	.71	.45	.38	.21	.30	.41	.20	.32	.37	.19	.31	.34	.21	.30
		LST	124	.81	.39	.22	.37	.33	.19	.38	.33	.17	.39	.34	.24	.40	.37
		TOT	124	1.52	.62	.42	.39	.43	.42	.39	.44	.38	.39	.44	.40	.41	.45
	A-4	MST	124	.59	.70	.38	.22	.31	.40	.21	.32	.35	.21	.31	.36	.23	.31
		LST	124	.80	.44	.29	.41	.39	.25	.42	.38	.21	.42	.39	.30	.44	.42
		TOT	124	1.39	.89	.44	.38	.44	.44	.37	.44	.38	.37	.43	.43	.40	.46

FORM B (Continued)

13	A-1	MST	124	10.52	12.78	-0.05	-0.18	-0.13	-0.05	-0.17	-0.13	.02	-0.16	-0.10	.00	-0.17	-0.11
		LST	124	19.87	23.28	.14	.24	.21	.12	.24	.21	.10	.25	.21	.15	.25	.23
		TOT	124	30.39	28.66	.09	.11	.11	.07	.12	.11	.09	.13	.13	.13	.13	.14
	A-2	MST	124	23.00	4.86	.04	.01	.02	.07	.01	.04	.09	.03	.06	.06	.01	.04
		LST	124	36.58	19.84	.18	.26	.25	.15	.27	.24	.13	.28	.25	.20	.28	.27
		TOT	124	59.58	21.48	.18	.24	.23	.16	.25	.23	.15	.26	.24	.20	.26	.26
	A-3	MST	124	.24	.43	-0.01	-0.11	-0.07	-0.01	-0.11	-0.07	.05	-0.09	-0.03	.02	-0.09	-0.05
		LST	124	.58	.49	.15	.25	.23	.13	.26	.23	.12	.27	.24	.17	.26	.25
		TOT	124	.82	.68	.10	.12	.12	.09	.12	.12	.12	.14	.15	.13	.13	.15
	A-4	MST	124	.21	.48	-0.03	-0.13	-0.09	-0.03	-0.13	-0.10	.03	-0.11	-0.06	.01	-0.11	-0.07
		LST	124	.52	.62	.14	.23	.21	.11	.23	.20	.09	.24	.20	.15	.25	.23
		TOT	124	.73	.85	.08	.09	.10	.06	.10	.09	.08	.11	.11	.11	.12	.13
14	A-1	MST	124	31.32	38.59	.19	.02	.10	.16	.03	.09	.16	.00	.08	.18	.01	.09
		LST	124	36.37	39.69	.36	.46	.45	.34	.48	.46	.26	.42	.41	.32	.41	.41
		TOT	124	67.69	59.54	.36	.32	.37	.33	.34	.37	.28	.28	.32	.33	.28	.33
	A-2	MST	124	39.11	25.44	.19	.02	.10	.17	.02	.09	.20	.01	.10	.20	.03	.11
		LST	124	43.78	31.33	.33	.45	.43	.31	.47	.45	.24	.41	.39	.29	.39	.39
		TOT	124	82.90	40.70	.38	.36	.40	.35	.38	.40	.31	.33	.37	.35	.32	.37
	A-3	MST	124	.55	.50	.19	.02	.10	.17	.02	.09	.20	.02	.11	.20	.03	.11
		LST	124	.53	.50	.31	.42	.40	.29	.44	.41	.21	.39	.36	.26	.36	.35
		TOT	124	1.08	.69	.36	.32	.36	.33	.34	.37	.30	.29	.34	.33	.29	.34
	A-4	MST	124	.48	.62	.19	.02	.10	.15	.03	.09	.17	.01	.09	.18	.02	.09
		LST	124	.50	.56	.34	.44	.43	.32	.46	.44	.25	.40	.39	.30	.38	.39
		TOT	124	.98	.88	.34	.29	.34	.31	.31	.34	.28	.26	.31	.32	.26	.31
15	A-1	MST	124	21.17	25.02	.13	.04	.09	.10	.04	.07	.10	.06	.09	.11	.05	.09
		LST	124	25.06	30.19	.36	.48	.46	.32	.47	.45	.34	.43	.45	.35	.46	.45
		TOT	124	46.23	43.26	.32	.36	.37	.28	.35	.36	.29	.34	.37	.31	.35	.37
	A-2	MST	124	28.94	11.13	.18	.09	.14	.17	.08	.13	.18	.08	.14	.19	.08	.14
		LST	124	37.75	25.23	.33	.46	.44	.30	.45	.43	.32	.42	.44	.33	.44	.43
		TOT	124	66.69	28.38	.36	.44	.44	.33	.43	.43	.36	.41	.44	.36	.43	.44
	A-3	MST	124	.47	.50	.14	.08	.11	.15	.07	.11	.17	.04	.11	.16	.06	.11
		LST	124	.45	.50	.28	.38	.37	.24	.39	.36	.28	.36	.38	.28	.38	.37
		TOT	124	.92	.69	.30	.33	.35	.28	.33	.34	.32	.29	.35	.32	.32	.35
	A-4	MST	124	.40	.62	.11	.04	.08	.10	.04	.07	.11	.03	.08	.11	.04	.08
		LST	124	.40	.58	.34	.42	.42	.29	.42	.40	.31	.38	.40	.33	.41	.42
		TOT	124	.80	.91	.29	.30	.32	.25	.29	.30	.27	.26	.31	.29	.28	.32

FORM B (Continued)

16	A-1	MST	122	14.52	21.86	.14	.03	.08	.14	.03	.08	.20	.03	.12	.17	.03	.10
		LST	123	24.17	43.87	.17	.36	.30	.18	.35	.30	.18	.35	.32	.18	.36	.31
		TOT	122	38.84	50.47	.21	.32	.30	.22	.32	.30	.24	.32	.33	.22	.33	.31
	A-2	MST	122	23.77	13.06	.11	.00	.05	.12	-0.00	.05	.19	-0.01	.09	.12	.00	.06
		LST	123	37.16	30.79	.15	.34	.28	.16	.33	.28	.16	.33	.29	.16	.34	.29
		TOT	122	61.12	33.31	.17	.31	.27	.19	.30	.28	.21	.30	.30	.19	.31	.28
	A-3	MST	122	.48	.50	.04	-0.04	-0.01	.06	-0.05	-0.00	.16	-0.04	.05	.06	-0.05	-0.00
		LST	123	.48	.50	.14	.32	.27	.15	.31	.27	.14	.31	.28	.15	.32	.27
		TOT	122	.96	.69	.12	.19	.18	.14	.18	.18	.21	.19	.23	.14	.19	.19
	A-4	MST	122	.45	.54	.08	-0.01	.03	.09	-0.01	.04	.18	-0.01	.09	.11	-0.01	.04
		LST	123	.24	.82	.16	.36	.30	.18	.35	.30	.18	.35	.32	.17	.36	.31
		TOT	122	.69	1.02	.17	.28	.26	.19	.27	.26	.24	.28	.30	.19	.29	.27
17	A-1	MST	124	18.34	22.76	.20	.16	.19	.20	.15	.19	.23	.12	.20	.24	.15	.21
		LST	124	44.21	42.36	.22	.50	.41	.20	.51	.42	.17	.45	.38	.21	.47	.40
		TOT	124	62.55	52.69	.26	.47	.41	.25	.47	.42	.24	.41	.39	.27	.44	.41
	A-2	MST	124	25.81	17.23	.15	.13	.15	.17	.12	.15	.20	.11	.17	.19	.14	.18
		LST	124	52.12	30.92	.20	.48	.39	.18	.50	.40	.14	.45	.37	.18	.46	.38
		TOT	124	77.94	36.73	.23	.46	.40	.23	.48	.41	.21	.43	.39	.24	.45	.40
	A-3	MST	124	.52	.50	.12	.10	.12	.15	.08	.12	.21	.07	.15	.17	.10	.15
		LST	124	.69	.46	.20	.48	.39	.19	.50	.40	.14	.46	.37	.18	.46	.38
		TOT	124	1.20	.70	.22	.39	.34	.23	.39	.35	.25	.36	.36	.25	.38	.36
	A-4	MST	124	.49	.55	.18	.15	.18	.20	.13	.18	.25	.12	.20	.23	.15	.21
		LST	124	.55	.72	.22	.51	.42	.21	.51	.42	.18	.45	.38	.22	.48	.41
		TOT	124	1.04	1.00	.26	.45	.40	.26	.44	.40	.26	.39	.39	.28	.43	.41
18	A-1	MST	123	39.91	31.65	.41	.36	.41	.40	.36	.41	.35	.35	.40	.40	.37	.42
		LST	123	24.40	31.13	.27	.35	.34	.26	.35	.34	.17	.33	.30	.25	.34	.33
		TOT	123	64.31	49.60	.43	.45	.48	.42	.45	.48	.33	.43	.45	.41	.45	.48
	A-2	MST	123	45.13	20.33	.34	.30	.34	.33	.30	.34	.30	.30	.35	.34	.32	.36
		LST	123	32.63	19.78	.21	.29	.27	.20	.29	.28	.13	.28	.25	.19	.27	.26
		TOT	123	77.76	30.79	.36	.38	.40	.35	.38	.41	.28	.38	.39	.35	.38	.41
	A-3	MST	123	.68	.47	.23	.16	.20	.20	.17	.20	.21	.19	.23	.25	.19	.24
		LST	123	.36	.48	.14	.20	.19	.15	.20	.20	.09	.22	.19	.14	.20	.19
		TOT	123	1.04	.70	.25	.24	.26	.24	.25	.27	.20	.27	.28	.26	.26	.29
	A-4	MST	123	.64	.56	.34	.27	.33	.31	.28	.33	.29	.29	.33	.36	.29	.36
		LST	123	.30	.57	.24	.29	.29	.24	.29	.29	.16	.29	.27	.23	.29	.29
		TOT	123	.94	.88	.37	.36	.40	.35	.37	.40	.29	.37	.39	.38	.38	.42

FORM B (Continued)

19	A-1	MST	124	25.88	32.59	.33	.28	.33	.30	.29	.32	.27	.30	.33	.31	.29	.33
		LST	124	34.17	26.35	.44	.52	.52	.43	.51	.52	.36	.49	.50	.41	.49	.50
		TOT	124	60.05	46.24	.48	.49	.53	.46	.49	.53	.40	.49	.52	.45	.48	.52
	A-2	MST	124	34.60	20.97	.31	.23	.28	.29	.24	.29	.30	.25	.31	.31	.25	.30
		LST	124	43.23	20.04	.38	.46	.46	.38	.45	.47	.30	.44	.44	.35	.44	.44
		TOT	124	77.83	30.90	.45	.45	.49	.45	.46	.50	.40	.46	.50	.44	.45	.49
	A-3	MST	124	.50	.50	.29	.19	.25	.29	.20	.26	.32	.21	.30	.31	.21	.28
		LST	124	.60	.49	.28	.32	.33	.29	.32	.34	.23	.33	.33	.26	.31	.32
		TOT	124	1.10	.73	.39	.35	.40	.39	.35	.41	.37	.36	.42	.39	.35	.41
	A-4	MST	124	.44	.60	.33	.25	.31	.31	.26	.31	.31	.28	.33	.32	.27	.32
		LST	124	.57	.56	.34	.39	.40	.35	.38	.41	.30	.38	.39	.33	.37	.39
		TOT	124	1.02	.90	.43	.41	.45	.42	.41	.46	.39	.42	.47	.42	.41	.46
20	A-1	MST	124	18.93	31.16	.24	.13	.19	.24	.10	.18	.30	.09	.21	.25	.12	.19
		LST	123	22.75	21.84	.30	.40	.39	.30	.40	.39	.31	.39	.41	.32	.41	.41
		TOT	123	41.61	38.60	.36	.33	.37	.36	.31	.37	.42	.29	.40	.39	.33	.39
	A-2	MST	124	30.95	18.74	.23	.12	.19	.24	.11	.18	.29	.08	.20	.24	.10	.17
		LST	123	32.20	15.50	.28	.33	.34	.28	.33	.34	.29	.33	.36	.30	.35	.36
		TOT	123	63.15	23.93	.36	.31	.36	.37	.30	.37	.42	.28	.39	.38	.30	.37
	A-3	MST	124	.66	.47	.25	.09	.17	.26	.06	.16	.35	.06	.21	.29	.09	.19
		LST	123	.77	.42	.28	.35	.35	.28	.35	.35	.29	.36	.38	.32	.37	.38
		TOT	123	1.43	.65	.37	.29	.35	.37	.27	.34	.44	.27	.39	.42	.30	.39
	A-4	MST	124	.46	.81	.22	.10	.17	.23	.06	.15	.30	.06	.19	.26	.10	.19
		LST	123	.67	.66	.28	.39	.38	.28	.39	.37	.29	.38	.40	.32	.41	.41
		TOT	123	1.12	1.07	.35	.32	.36	.34	.29	.34	.41	.28	.39	.40	.33	.39
21	A-1	MST	124	37.72	20.06	.23	.19	.22	.19	.20	.21	.13	.14	.16	.18	.16	.19
		LST	124	40.13	31.56	.16	.27	.24	.15	.30	.26	.15	.29	.26	.15	.28	.25
		TOT	124	77.85	40.75	.24	.31	.30	.21	.33	.30	.18	.29	.28	.21	.29	.28
	A-2	MST	124	41.82	12.76	.22	.22	.24	.19	.22	.22	.14	.14	.16	.18	.18	.20
		LST	124	44.97	25.72	.12	.22	.19	.10	.25	.21	.11	.23	.21	.12	.22	.19
		TOT	124	86.79	31.16	.19	.27	.25	.16	.29	.26	.15	.25	.24	.17	.25	.24
	A-3	MST	124	.90	.30	.22	.22	.24	.19	.22	.23	.14	.14	.16	.18	.18	.20
		LST	124	.59	.49	.06	.17	.13	.05	.20	.15	.08	.17	.15	.07	.17	.14
		TOT	124	1.49	.63	.15	.23	.22	.13	.26	.22	.12	.20	.20	.14	.21	.20
	A-4	MST	124	.89	.36	.19	.16	.19	.15	.16	.17	.10	.11	.12	.15	.13	.15
		LST	124	.53	.60	.13	.25	.21	.12	.28	.23	.13	.26	.24	.12	.26	.22
		TOT	124	1.42	.76	.19	.27	.26	.17	.29	.26	.15	.26	.24	.17	.26	.24

FORM B (Continued)

22	A-1	MST	124	31.52	33.93	.41	.26	.35	.39	.26	.35	.34	.24	.32	.42	.22	.34
		LST	123	19.94	34.23	.38	.46	.46	.37	.45	.46	.36	.45	.48	.38	.48	.48
		TOT	123	51.50	54.18	.50	.45	.51	.48	.45	.51	.44	.44	.51	.50	.44	.52
	A-2	MST	124	38.57	25.23	.35	.17	.27	.35	.18	.27	.32	.15	.26	.35	.14	.25
		LST	123	30.59	16.55	.31	.45	.42	.30	.44	.42	.29	.46	.44	.30	.48	.44
		TOT	123	69.24	32.43	.43	.37	.43	.42	.37	.43	.39	.36	.43	.43	.35	.42
	A-3	MST	124	.42	.49	.24	.04	.14	.25	.06	.15	.25	.03	.14	.24	.00	.12
		LST	123	.71	.45	.24	.41	.36	.23	.40	.36	.22	.44	.40	.23	.44	.39
		TOT	123	1.13	.71	.32	.29	.33	.32	.30	.34	.31	.31	.36	.32	.29	.33
	A-4	MST	124	.32	.64	.35	.18	.27	.34	.19	.28	.30	.16	.26	.37	.14	.26
		LST	123	.61	.66	.30	.38	.38	.29	.38	.37	.28	.41	.41	.30	.42	.40
		TOT	123	.93	1.03	.41	.36	.41	.40	.36	.41	.37	.36	.42	.42	.36	.43
23	A-1	MST	124	18.60	29.20	.32	.29	.33	.32	.29	.33	.26	.29	.32	.31	.29	.33
		LST	121	13.35	34.21	.15	.39	.31	.14	.39	.31	.11	.42	.33	.14	.40	.32
		TOT	121	32.28	50.57	.28	.44	.40	.28	.44	.40	.23	.46	.41	.27	.45	.41
	A-2	MST	124	31.98	18.74	.32	.31	.34	.34	.31	.35	.29	.31	.35	.31	.30	.33
		LST	121	28.11	19.23	.14	.36	.28	.12	.37	.29	.10	.40	.31	.13	.38	.30
		TOT	121	60.37	29.88	.29	.43	.40	.29	.44	.41	.24	.46	.42	.28	.44	.41
	A-3	MST	124	.48	.50	.31	.32	.34	.34	.31	.36	.32	.29	.35	.30	.29	.33
		LST	121	.40	.49	.12	.33	.26	.10	.35	.26	.07	.38	.29	.12	.35	.28
		TOT	121	.89	.75	.29	.43	.40	.29	.44	.41	.26	.45	.42	.28	.43	.40
	A-4	MST	124	.40	.63	.34	.31	.35	.34	.30	.35	.31	.29	.34	.34	.30	.35
		LST	121	.29	.66	.13	.41	.31	.12	.41	.31	.10	.42	.32	.12	.41	.32
		TOT	121	.69	1.02	.30	.46	.42	.29	.45	.42	.25	.46	.42	.29	.46	.43
24	A-1	MST	124	9.86	13.03	.28	.20	.25	.25	.20	.24	.21	.23	.25	.26	.21	.26
		LST	123	7.83	20.72	.29	.27	.30	.27	.26	.29	.29	.20	.27	.30	.22	.28
		TOT	123	17.64	26.93	.36	.31	.36	.33	.30	.35	.32	.27	.34	.36	.28	.35
	A-2	MST	124	25.57	7.24	.30	.21	.27	.27	.21	.26	.23	.24	.27	.29	.23	.28
		LST	123	24.37	11.15	.31	.30	.32	.28	.29	.31	.31	.24	.31	.33	.26	.32
		TOT	123	49.93	14.75	.38	.34	.38	.35	.33	.37	.35	.31	.37	.40	.32	.39
	A-3	MST	124	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		LST	123	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		TOT	123	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
	A-4	MST	124	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		LST	123	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		TOT	123	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00

FORM B (Continued)

25	A-1	MST	124	27.81	34.26	.27	.14	.21	.25	.16	.21	.24	.16	.22	.28	.14	.22
		LST	124	28.88	34.61	.22	.32	.30	.22	.32	.31	.19	.30	.29	.23	.30	.30
		TOT	124	56.69	51.76	.33	.30	.34	.31	.32	.35	.29	.30	.34	.33	.29	.34
	A-2	MST	124	34.72	23.17	.24	.16	.21	.24	.17	.22	.23	.16	.22	.25	.15	.21
		LST	124	37.43	24.49	.20	.28	.26	.19	.29	.27	.16	.26	.25	.21	.26	.26
		TOT	124	72.15	35.55	.30	.29	.32	.29	.31	.33	.26	.28	.32	.30	.28	.32
	A-3	MST	124	.44	.50	.20	.13	.18	.20	.15	.19	.19	.13	.18	.20	.13	.18
		LST	124	.44	.50	.18	.20	.21	.17	.21	.21	.15	.19	.20	.21	.18	.22
		TOT	124	.88	.75	.25	.22	.25	.25	.24	.27	.23	.21	.25	.27	.21	.26
	A-4	MST	124	.40	.58	.25	.13	.20	.23	.15	.20	.22	.15	.21	.26	.13	.21
		LST	124	.39	.58	.21	.26	.26	.21	.27	.27	.19	.25	.26	.23	.26	.27
		TOT	124	.78	.89	.30	.25	.30	.29	.27	.31	.26	.26	.30	.32	.25	.31
26	A-1	MST	124	13.65	23.30	.27	.21	.26	.27	.18	.24	.21	.16	.21	.26	.19	.24
		LST	124	11.69	23.43	.32	.40	.40	.30	.38	.38	.23	.37	.36	.28	.40	.38
		TOT	124	25.34	37.12	.37	.39	.41	.36	.35	.39	.27	.34	.36	.34	.37	.39
	A-2	MST	124	28.05	16.13	.24	.15	.20	.25	.13	.20	.19	.12	.17	.22	.13	.19
		LST	124	26.19	13.64	.34	.45	.43	.33	.43	.42	.25	.42	.40	.29	.43	.41
		TOT	124	54.23	23.32	.36	.37	.39	.36	.34	.38	.28	.33	.35	.33	.34	.37
	A-3	MST	124	.44	.50	.19	.10	.15	.21	.08	.15	.18	.08	.14	.20	.10	.15
		LST	124	.63	.48	.34	.44	.43	.33	.43	.42	.25	.42	.40	.28	.43	.41
		TOT	124	1.07	.75	.34	.35	.38	.35	.33	.37	.28	.33	.35	.31	.34	.36
	A-4	MST	124	.27	.74	.27	.19	.24	.27	.15	.22	.21	.15	.20	.26	.17	.23
		LST	124	.53	.67	.30	.39	.38	.29	.37	.37	.22	.37	.35	.26	.39	.37
		TOT	124	.80	1.12	.36	.36	.39	.35	.32	.37	.27	.31	.34	.33	.34	.37
27	A-1	MST	124	20.06	32.76	.44	.35	.42	.42	.34	.41	.38	.32	.40	.41	.33	.40
		LST	122	17.64	28.88	.11	.36	.27	.09	.36	.27	.04	.35	.25	.08	.36	.26
		TOT	122	38.38	48.89	.36	.45	.45	.33	.45	.44	.27	.42	.41	.32	.44	.43
	A-2	MST	124	31.81	20.32	.44	.35	.42	.43	.34	.42	.38	.32	.40	.39	.32	.39
		LST	122	29.39	18.91	.12	.31	.25	.09	.31	.24	.04	.33	.24	.10	.33	.26
		TOT	122	61.72	29.63	.38	.44	.45	.35	.44	.44	.27	.44	.42	.33	.44	.43
	A-3	MST	124	.44	.50	.29	.20	.26	.27	.20	.25	.28	.17	.25	.27	.18	.24
		LST	122	.39	.49	.14	.27	.23	.10	.28	.22	.05	.31	.23	.12	.30	.25
		TOT	122	.84	.75	.28	.31	.32	.24	.32	.31	.21	.32	.31	.26	.32	.32
	A-4	MST	124	.35	.64	.34	.26	.32	.31	.25	.30	.31	.22	.30	.33	.24	.31
		LST	122	.32	.60	.14	.34	.27	.10	.34	.26	.06	.34	.26	.12	.35	.27
		TOT	122	.68	.98	.31	.38	.38	.26	.38	.36	.24	.36	.36	.29	.38	.38

FORM B (Continued)

28	A-1	MST	124	21.46	22.61	.12	.05	.08	.12	.04	.08	.16	.04	.11	.15	.07	.11
		LST	124	12.48	29.51	.16	.21	.21	.19	.19	.21	.15	.23	.22	.16	.24	.22
		TOT	124	33.94	41.08	.18	.18	.19	.21	.16	.19	.19	.19	.22	.19	.21	.22
	A-2	MST	124	33.44	15.05	.13	.03	.08	.14	.02	.08	.20	.02	.11	.18	.05	.12
		LST	124	27.65	15.38	.17	.26	.24	.20	.25	.25	.15	.28	.26	.15	.28	.25
		TOT	124	61.08	22.37	.20	.20	.22	.24	.18	.23	.23	.20	.25	.22	.23	.25
	A-3	MST	124	.53	.50	.10	-0.02	.03	.11	-0.03	.03	.19	-0.04	.07	.15	-0.01	.07
		LST	124	.43	.49	.15	.24	.21	.18	.23	.23	.12	.28	.24	.12	.26	.22
		TOT	124	.96	.66	.18	.16	.18	.22	.15	.19	.23	.18	.23	.21	.19	.22
	A-4	MST	124	.49	.57	.09	.01	.05	.10	.00	.05	.15	.00	.07	.14	.03	.09
		LST	124	.27	.71	.12	.18	.17	.15	.16	.17	.12	.21	.20	.12	.21	.19
		TOT	124	.77	1.00	.14	.14	.15	.17	.12	.15	.17	.15	.18	.17	.17	.19
29	A-1	MST	124	16.06	31.36	.23	.15	.20	.21	.13	.18	.19	.13	.18	.23	.14	.20
		LST	124	13.15	24.28	.32	.20	.27	.31	.19	.26	.26	.18	.25	.29	.18	.25
		TOT	124	29.22	43.25	.34	.22	.30	.32	.20	.28	.28	.20	.27	.33	.21	.29
	A-2	MST	124	30.90	20.18	.21	.15	.19	.20	.13	.18	.19	.13	.18	.22	.14	.19
		LST	124	26.84	14.89	.21	.09	.15	.19	.10	.15	.17	.13	.17	.18	.10	.15
		TOT	124	57.74	25.47	.29	.17	.24	.27	.16	.23	.25	.18	.24	.28	.17	.24
	A-3	MST	124	.44	.50	.18	.14	.17	.17	.12	.16	.17	.12	.16	.18	.13	.17
		LST	124	.40	.49	.15	.02	.08	.14	.03	.08	.12	.08	.12	.12	.04	.08
		TOT	124	.85	.67	.24	.12	.18	.23	.11	.18	.21	.15	.20	.23	.13	.19
	A-4	MST	124	.34	.66	.23	.14	.20	.21	.12	.18	.19	.13	.18	.24	.14	.21
		LST	124	.32	.62	.28	.15	.22	.27	.15	.22	.23	.17	.22	.26	.15	.22
		TOT	124	.66	.97	.34	.19	.28	.31	.18	.26	.28	.20	.27	.33	.19	.28
30	A-1	MST	124	30.45	27.96	.38	.40	.43	.36	.40	.42	.30	.39	.40	.35	.39	.41
		LST	124	31.69	25.30	.26	.34	.33	.27	.34	.34	.21	.33	.32	.22	.32	.31
		TOT	124	62.14	39.18	.44	.51	.52	.43	.51	.52	.35	.49	.49	.39	.49	.49
	A-2	MST	124	37.54	20.08	.35	.37	.39	.34	.37	.39	.29	.35	.38	.33	.36	.38
		LST	124	35.10	18.56	.24	.31	.31	.24	.32	.31	.19	.30	.29	.21	.30	.29
		TOT	124	72.65	27.88	.41	.48	.49	.40	.47	.49	.33	.45	.46	.38	.46	.47
	A-3	MST	124	.37	.48	.15	.21	.20	.16	.20	.21	.18	.19	.21	.17	.20	.20
		LST	124	.46	.50	.21	.28	.27	.20	.28	.27	.16	.26	.25	.19	.27	.26
		TOT	124	.83	.67	.26	.36	.35	.27	.36	.35	.25	.33	.34	.26	.34	.34
	A-4	MST	124	.35	.51	.17	.27	.25	.19	.26	.25	.20	.25	.26	.19	.25	.25
		LST	124	.44	.54	.23	.33	.31	.23	.33	.32	.18	.31	.30	.20	.31	.29
		TOT	124	.79	.74	.29	.42	.39	.30	.42	.40	.27	.40	.40	.28	.40	.38

FORM B (Continued)

31	A-1	MST	124	12.11	16.65	.24	.16	.21	.19	.16	.19	.22	.09	.17	.28	.14	.22
		LST	124	19.65	29.07	.25	.26	.28	.28	.27	.30	.26	.29	.32	.23	.29	.29
		TOT	124	31.76	32.83	.34	.31	.35	.35	.32	.37	.34	.30	.37	.35	.33	.37
	A-2	MST	124	23.51	7.80	.13	.06	.09	.09	.07	.08	.15	.00	.08	.18	.05	.11
		LST	124	34.98	22.92	.25	.27	.28	.28	.28	.31	.26	.30	.32	.23	.29	.29
		TOT	124	58.48	23.64	.29	.28	.31	.30	.29	.33	.30	.29	.34	.28	.30	.32
	A-3	MST	124	.31	.46	.06	-0.01	.02	.03	.00	.01	.11	-0.05	.02	.11	-0.01	.04
		LST	124	.52	.50	.25	.27	.28	.27	.29	.31	.25	.30	.32	.22	.29	.29
		TOT	124	.82	.65	.23	.20	.23	.23	.22	.25	.27	.19	.26	.25	.22	.25
	A-4	MST	124	.27	.51	.17	.09	.14	.13	.10	.12	.18	.03	.11	.22	.08	.15
		LST	124	.44	.63	.25	.24	.27	.28	.25	.29	.26	.26	.30	.23	.28	.29
		TOT	124	.72	.80	.30	.25	.30	.30	.26	.31	.32	.23	.31	.33	.27	.33
32	A-1	MST	123	21.57	34.03	.26	.19	.24	.26	.20	.25	.27	.19	.26	.26	.18	.24
		LST	123	19.98	31.61	.19	.35	.31	.19	.33	.30	.17	.35	.32	.19	.36	.32
		TOT	123	41.54	51.28	.29	.35	.35	.29	.34	.35	.28	.34	.37	.29	.34	.35
	A-2	MST	123	32.98	24.52	.26	.20	.24	.27	.20	.26	.27	.19	.26	.24	.19	.23
		LST	123	33.32	22.32	.23	.36	.33	.22	.35	.32	.20	.38	.35	.23	.37	.35
		TOT	123	66.29	35.76	.32	.36	.37	.32	.35	.38	.31	.37	.40	.31	.36	.38
	A-3	MST	123	.44	.50	.25	.19	.24	.26	.19	.25	.25	.19	.25	.23	.19	.23
		LST	123	.49	.50	.23	.35	.33	.23	.34	.32	.21	.37	.35	.24	.37	.35
		TOT	123	.79	.97	.30	.35	.35	.29	.34	.35	.29	.35	.37	.30	.35	.36
	A-4	MST	123	.37	.60	.25	.19	.23	.25	.19	.24	.26	.18	.25	.25	.17	.23
		LST	123	.41	.62	.22	.36	.33	.22	.35	.32	.20	.37	.34	.22	.38	.34
		TOT	123	.93	.76	.32	.36	.37	.32	.35	.37	.31	.37	.40	.31	.37	.38
33	A-1	MST	124	36.06	45.39	.42	.31	.39	.41	.31	.38	.32	.31	.36	.39	.30	.37
		LST	124	17.27	29.42	.38	.33	.38	.37	.32	.37	.29	.34	.37	.36	.33	.38
		TOT	124	53.32	62.01	.49	.39	.47	.47	.38	.46	.37	.39	.44	.45	.38	.45
	A-2	MST	124	45.98	33.54	.42	.30	.38	.40	.29	.37	.32	.28	.35	.38	.28	.36
		LST	124	28.13	17.77	.27	.22	.27	.26	.24	.27	.19	.28	.28	.25	.24	.27
		TOT	124	74.11	42.02	.45	.33	.41	.43	.33	.41	.34	.35	.39	.41	.33	.40
	A-3	MST	124	.59	.49	.41	.30	.38	.40	.28	.37	.32	.28	.34	.38	.28	.36
		LST	124	.35	.48	.18	.15	.17	.16	.17	.18	.11	.22	.20	.15	.17	.18
		TOT	124	.94	.75	.38	.29	.35	.36	.29	.35	.28	.32	.35	.35	.29	.35
	A-4	MST	124	.50	.65	.42	.31	.38	.40	.30	.38	.32	.31	.36	.39	.30	.38
		LST	124	.29	.58	.30	.26	.30	.29	.27	.30	.22	.31	.31	.28	.28	.31
		TOT	124	.79	1.00	.45	.35	.43	.43	.35	.42	.33	.38	.41	.42	.36	.42

FORM B (Continued)

34	A-1	MST	123	18.68	30.86	.39	.26	.34	.37	.25	.33	.31	.21	.29	.35	.22	.31
		LST	121	12.72	27.12	.18	.22	.22	.18	.21	.22	.15	.26	.24	.16	.25	.23
		TOT	121	31.67	42.53	.40	.33	.39	.38	.32	.38	.31	.32	.36	.36	.32	.37
	A-2	MST	123	32.46	18.60	.40	.27	.35	.38	.26	.34	.32	.22	.31	.35	.23	.31
		LST	121	26.49	16.23	.17	.19	.20	.17	.20	.20	.12	.23	.22	.14	.21	.20
		TOT	121	59.23	25.44	.40	.33	.39	.38	.32	.38	.31	.32	.36	.35	.31	.36
	A-3	MST	123	.46	.50	.37	.25	.32	.35	.25	.32	.30	.21	.29	.32	.21	.28
		LST	121	.32	.47	.08	.10	.10	.07	.11	.10	.06	.15	.13	.08	.13	.12
		TOT	121	.79	.62	.36	.28	.34	.34	.28	.34	.29	.29	.33	.32	.28	.32
	A-4	MST	123	.41	.60	.43	.29	.37	.40	.28	.36	.34	.24	.33	.39	.26	.34
		LST	121	.25	.58	.15	.19	.19	.15	.18	.18	.13	.23	.22	.14	.22	.21
		TOT	121	.66	.84	.41	.34	.40	.39	.33	.39	.33	.34	.38	.37	.34	.39
35	A-1	MST	124	19.59	27.05	.38	.34	.39	.39	.32	.39	.39	.32	.40	.40	.33	.40
		LST	124	16.35	29.69	.09	.27	.21	.08	.25	.20	.02	.24	.17	.06	.27	.20
		TOT	124	35.94	43.88	.29	.39	.38	.30	.37	.37	.25	.36	.36	.29	.39	.38
	A-2	MST	124	32.19	16.56	.39	.36	.40	.39	.34	.40	.40	.35	.43	.41	.36	.42
		LST	124	28.36	17.72	.04	.23	.16	.04	.21	.15	-0.02	.22	.13	.01	.23	.15
		TOT	124	60.55	26.00	.27	.38	.36	.28	.36	.36	.24	.37	.36	.26	.39	.37
	A-3	MST	124	.69	.46	.39	.34	.39	.40	.32	.39	.42	.32	.42	.42	.34	.41
		LST	124	.37	.48	-0.01	.17	.10	-0.00	.15	.10	-0.06	.17	.09	-0.05	.18	.09
		TOT	124	1.06	.71	.25	.33	.32	.26	.31	.32	.23	.33	.33	.24	.34	.33
	A-4	MST	124	.62	.60	.37	.33	.38	.38	.31	.38	.38	.31	.39	.40	.32	.39
		LST	124	.19	.72	.07	.26	.20	.07	.24	.19	.01	.24	.16	.05	.26	.19
		TOT	124	.81	1.03	.27	.38	.36	.28	.35	.35	.23	.35	.34	.27	.37	.36
36	A-1	MST	124	12.64	21.32	.30	.36	.36	.27	.37	.36	.23	.32	.33	.28	.34	.35
		LST	124	7.02	23.43	.09	.16	.14	.11	.15	.14	.02	.22	.16	.03	.20	.14
		TOT	124	19.66	34.81	.25	.33	.32	.24	.33	.32	.15	.35	.31	.19	.34	.31
	A-2	MST	124	26.39	10.23	.33	.38	.39	.30	.39	.39	.24	.35	.35	.30	.36	.37
		LST	124	23.57	11.95	.10	.17	.15	.11	.16	.15	.03	.23	.17	.04	.21	.15
		TOT	124	49.96	17.44	.26	.34	.33	.25	.34	.33	.16	.36	.32	.20	.36	.32
	A-3	MST	124	.36	.48	.12	.13	.14	.11	.14	.14	.18	.10	.16	.13	.10	.13
		LST	124	.53	.50	.09	.16	.14	.10	.15	.14	.01	.22	.15	.02	.18	.13
		TOT	124	.90	.68	.15	.21	.20	.15	.21	.20	.13	.23	.22	.11	.21	.18
	A-4	MST	124	.18	.72	.27	.30	.31	.24	.31	.31	.25	.27	.30	.26	.28	.30
		LST	124	.44	.65	.06	.15	.12	.06	.15	.12	-0.04	.21	.12	-0.02	.18	.11
		TOT	124	.62	1.07	.21	.29	.28	.20	.30	.28	.14	.31	.27	.16	.30	.26

FORM B (Continued)

37	A-1	MST	124	17.73	27.24	.27	.10	.19	.30	.10	.20	.32	.08	.21	.25	.08	.17
		LST	123	7.48	17.52	.32	.44	.42	.36	.43	.44	.31	.43	.44	.30	.44	.42
		TOT	123	25.46	36.38	.35	.29	.34	.40	.29	.37	.38	.28	.37	.33	.28	.33
	A-2	MST	124	32.41	19.29	.26	.09	.18	.30	.09	.19	.31	.06	.19	.24	.07	.16
		LST	123	22.60	7.40	.27	.38	.36	.29	.39	.38	.25	.38	.38	.24	.38	.35
		TOT	123	55.14	22.53	.31	.21	.27	.35	.21	.29	.34	.18	.29	.28	.18	.25
	A-3	MST	124	.48	.50	.24	.07	.16	.28	.07	.17	.31	.04	.18	.23	.05	.14
		LST	123	.24	.42	.16	.15	.17	.20	.14	.18	.13	.20	.20	.16	.17	.18
		TOT	123	.72	.68	.28	.16	.22	.32	.14	.24	.30	.16	.25	.27	.15	.22
	A-4	MST	124	.42	.61	.24	.09	.17	.27	.09	.18	.28	.08	.19	.22	.07	.15
		LST	123	.16	.53	.28	.33	.33	.32	.31	.35	.25	.35	.35	.27	.34	.34
		TOT	123	.59	.88	.33	.26	.32	.38	.25	.34	.34	.27	.34	.31	.26	.31
38	A-1	MST	124	11.61	25.25	.18	.10	.15	.17	.11	.15	.19	.13	.18	.18	.11	.16
		LST	123	10.29	25.48	.16	.16	.17	.17	.14	.17	.15	.11	.15	.14	.14	.15
		TOT	123	21.69	37.83	.22	.18	.21	.23	.17	.21	.22	.16	.22	.21	.17	.21
	A-2	MST	124	25.58	13.96	.17	.10	.14	.17	.11	.15	.17	.13	.17	.17	.11	.15
		LST	123	24.98	12.69	.15	.18	.18	.17	.15	.18	.16	.13	.17	.13	.16	.16
		TOT	123	50.45	19.70	.22	.19	.22	.22	.18	.22	.22	.18	.23	.20	.18	.21
	A-3	MST	124	.36	.48	.19	.10	.15	.18	.11	.16	.23	.11	.19	.22	.10	.17
		LST	123	.31	.46	.01	.08	.06	.04	.06	.05	.03	.07	.06	-0.02	.08	.04
		TOT	123	.67	.63	.15	.14	.15	.16	.13	.15	.19	.13	.18	.15	.14	.16
	A-4	MST	124	.31	.57	.23	.13	.19	.22	.14	.19	.25	.13	.21	.26	.13	.21
		LST	123	.22	.59	.16	.17	.18	.18	.14	.17	.17	.13	.17	.14	.16	.16
		TOT	123	.52	.85	.26	.21	.25	.27	.19	.24	.28	.17	.25	.26	.20	.25
39	A-1	MST	124	7.79	14.32	.18	.17	.19	.15	.19	.19	.08	.15	.14	.09	.16	.14
		LST	120	4.73	14.05	.19	.28	.26	.18	.28	.26	.10	.25	.21	.15	.25	.23
		TOT	120	12.72	23.09	.23	.28	.28	.21	.29	.28	.11	.25	.22	.15	.26	.23
	A-2	MST	124	23.60	7.22	.28	.24	.27	.29	.23	.28	.17	.19	.21	.17	.19	.20
		LST	120	25.92	11.73	.28	.40	.38	.25	.42	.38	.16	.37	.33	.24	.38	.35
		TOT	120	49.53	14.76	.37	.43	.44	.35	.45	.45	.22	.39	.37	.28	.40	.39
	A-3	MST	124	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		LST	120	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		TOT	120	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
	A-4	MST	124	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		LST	120	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		TOT	120	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00

FORM B (Continued)

40	A-1	MST	123	20.67	27.42	.14	.10	.13	.12	.08	.11	.11	.10	.12	.15	.10	.14
		LST	124	17.40	31.83	.10	.22	.18	.10	.20	.17	.06	.20	.16	.09	.24	.19
		TOT	123	37.80	43.74	.17	.22	.21	.17	.20	.20	.13	.20	.20	.17	.24	.23
	A-2	MST	123	30.58	16.71	.11	.08	.10	.10	.07	.09	.11	.09	.11	.13	.09	.11
		LST	124	28.44	17.94	.08	.20	.16	.09	.18	.15	.05	.19	.15	.08	.23	.18
		TOT	123	58.82	26.22	.14	.19	.18	.14	.17	.17	.12	.19	.18	.15	.21	.20
	A-3	MST	123	.64	.48	.07	.03	.05	.07	.02	.05	.11	.03	.07	.11	.03	.07
		LST	124	.27	.45	.02	.09	.06	.02	.07	.05	-0.00	.11	.07	.04	.12	.09
		TOT	123	.91	.72	.07	.08	.08	.08	.06	.07	.08	.09	.10	.11	.09	.11
	A-4	MST	123	.59	.60	.10	.03	.07	.09	.02	.06	.11	.04	.08	.14	.04	.09
		LST	124	.06	.69	.06	.15	.12	.07	.13	.11	.04	.14	.11	.07	.18	.15
		TOT	123	.64	.96	.12	.13	.13	.12	.11	.13	.11	.13	.14	.14	.15	.16
41	A-1	MST	123	38.86	29.19	.29	.05	.17	.29	.05	.17	.34	.03	.18	.31	.05	.18
		LST	123	35.55	44.04	.39	.49	.48	.35	.49	.48	.32	.44	.45	.37	.46	.47
		TOT	123	74.41	57.50	.44	.40	.45	.42	.40	.45	.42	.35	.43	.44	.38	.45
	A-2	MST	123	43.81	22.25	.25	.05	.15	.27	.04	.15	.30	.02	.16	.26	.05	.15
		LST	123	44.71	34.32	.37	.51	.49	.34	.52	.49	.30	.46	.46	.35	.49	.48
		TOT	123	88.52	42.38	.43	.44	.47	.42	.44	.48	.40	.38	.45	.42	.42	.47
	A-3	MST	123	.65	.48	.22	.05	.13	.25	.04	.14	.27	.02	.14	.23	.05	.14
		LST	123	.54	.50	.35	.51	.48	.32	.52	.48	.29	.46	.45	.33	.49	.47
		TOT	123	1.20	.71	.40	.39	.43	.40	.39	.43	.38	.34	.41	.39	.38	.42
	A-4	MST	123	.63	.53	.26	.05	.15	.27	.05	.16	.31	.03	.17	.28	.05	.16
		LST	123	.49	.60	.37	.47	.46	.33	.47	.45	.30	.42	.42	.36	.44	.45
		TOT	123	1.11	.88	.41	.35	.41	.39	.35	.41	.39	.30	.39	.42	.33	.41
42	A-1	MST	124	61.80	42.16	.42	.32	.39	.46	.30	.40	.35	.28	.36	.37	.31	.37
		LST	124	33.39	31.19	.21	.41	.35	.18	.41	.34	.13	.39	.32	.19	.39	.33
		TOT	124	95.19	58.18	.42	.45	.47	.43	.44	.48	.32	.41	.43	.37	.43	.45
	A-2	MST	124	65.55	34.48	.39	.29	.36	.43	.27	.37	.32	.25	.33	.35	.28	.34
		LST	124	37.02	26.22	.24	.46	.40	.22	.47	.40	.14	.45	.37	.20	.44	.38
		TOT	124	102.56	47.74	.42	.46	.48	.43	.45	.49	.31	.43	.44	.36	.45	.45
	A-3	MST	124	.76	.43	.38	.27	.35	.42	.25	.36	.31	.24	.31	.33	.27	.33
		LST	124	.49	.50	.24	.44	.38	.22	.46	.39	.14	.45	.37	.21	.43	.37
		TOT	124	1.25	.73	.39	.47	.47	.40	.47	.48	.28	.46	.44	.34	.46	.45
	A-4	MST	124	.73	.49	.43	.31	.39	.46	.29	.40	.36	.27	.36	.38	.30	.37
		LST	124	.48	.53	.22	.41	.36	.20	.43	.36	.14	.42	.34	.20	.41	.35
		TOT	124	1.21	.81	.41	.46	.48	.41	.46	.48	.31	.44	.44	.36	.45	.46

43	A-1	MST	124	13.84	27.28	.17	.12	.16	.16	.13	.15	.21	.11	.18	.22	.14	.19
		LST	124	6.32	33.60	.28	.32	.33	.24	.30	.30	.18	.32	.30	.26	.35	.35
		TOT	124	20.16	48.91	.29	.29	.31	.25	.27	.29	.24	.28	.31	.30	.32	.35
	A-2	MST	124	26.60	20.60	.15	.12	.14	.16	.12	.15	.20	.11	.17	.19	.14	.18
		LST	124	26.34	16.08	.38	.41	.43	.33	.39	.40	.24	.40	.39	.34	.41	.42
		TOT	124	52.94	27.52	.34	.33	.36	.31	.32	.34	.29	.32	.36	.34	.35	.38
	A-3	MST	124	.32	.47	.13	.11	.13	.14	.10	.13	.19	.10	.16	.17	.13	.16
		LST	124	.60	.49	.34	.37	.39	.29	.35	.36	.21	.38	.36	.30	.39	.39
		TOT	124	.92	.73	.31	.32	.34	.29	.30	.33	.26	.32	.34	.31	.35	.37
	A-4	MST	124	.14	.70	.16	.11	.14	.14	.12	.14	.19	.10	.16	.21	.14	.18
		LST	124	.36	.84	.25	.29	.29	.21	.26	.26	.16	.30	.27	.23	.32	.31
		TOT	124	.50	1.24	.25	.26	.28	.22	.24	.26	.21	.26	.27	.27	.30	.31
44	A-1	MST	124	13.44	22.04	.13	.14	.15	.15	.12	.14	.20	.10	.17	.15	.14	.16
		LST	124	25.73	31.66	.37	.38	.41	.37	.36	.40	.37	.36	.42	.39	.39	.43
		TOT	124	39.17	44.63	.33	.34	.36	.34	.31	.36	.36	.31	.38	.35	.35	.39
	A-2	MST	124	25.56	13.34	.15	.18	.18	.16	.16	.18	.20	.15	.20	.15	.19	.19
		LST	124	36.22	20.84	.38	.38	.41	.37	.37	.40	.36	.37	.42	.39	.39	.43
		TOT	124	61.77	28.15	.35	.37	.39	.35	.35	.38	.36	.34	.40	.36	.38	.41
	A-3	MST	124	.23	.42	.09	.09	.09	.10	.07	.09	.13	.07	.11	.07	.09	.09
		LST	124	.54	.50	.36	.35	.39	.35	.34	.38	.34	.34	.39	.38	.36	.40
		TOT	124	.77	.73	.30	.29	.32	.29	.28	.31	.31	.27	.33	.30	.30	.33
	A-4	MST	124	.19	.49	.11	.09	.11	.13	.07	.10	.18	.06	.13	.12	.10	.12
		LST	124	.48	.62	.38	.39	.42	.38	.37	.41	.39	.37	.43	.40	.39	.44
		TOT	124	.67	.90	.32	.31	.35	.33	.29	.34	.36	.28	.37	.34	.32	.37
45	A-1	MST	124	8.68	25.54	.11	.17	.16	.09	.18	.15	.07	.17	.14	.12	.18	.17
		LST	123	8.60	24.04	.13	.25	.22	.08	.26	.20	.02	.28	.20	.12	.27	.23
		TOT	123	17.59	34.74	.17	.31	.27	.13	.31	.26	.08	.32	.25	.18	.32	.29
	A-2	MST	124	24.68	16.57	.09	.16	.14	.08	.16	.14	.07	.15	.13	.11	.17	.16
		LST	123	23.83	11.65	.09	.24	.19	.04	.25	.18	-0.02	.27	.17	.08	.26	.20
		TOT	123	48.66	19.58	.14	.28	.24	.10	.29	.23	.06	.29	.22	.14	.30	.26
	A-3	MST	124	.29	.45	.04	.12	.10	.04	.12	.10	.05	.12	.10	.06	.13	.11
		LST	123	.57	.50	.06	.22	.16	.01	.22	.15	-0.06	.27	.15	.03	.25	.17
		TOT	123	.86	.64	.08	.25	.19	.04	.26	.18	-0.00	.29	.19	.07	.29	.22
	A-4	MST	124	.08	.70	.07	.17	.14	.05	.17	.13	.03	.16	.12	.09	.18	.16
		LST	123	.46	.68	.12	.24	.20	.07	.24	.19	.01	.28	.19	.10	.27	.22
		TOT	123	.55	.99	.14	.29	.24	.09	.29	.23	.04	.32	.23	.14	.32	.27

FORM B (Continued)

46	A-1	MST	124	28.72	39.95	.56	.54	.60	.53	.53	.58	.42	.49	.53	.52	.52	.57
		LST	124	49.05	48.71	.40	.58	.54	.38	.59	.55	.26	.52	.47	.30	.53	.48
		TOT	124	77.77	75.92	.55	.66	.66	.52	.65	.66	.39	.59	.58	.47	.61	.61
	A-2	MST	124	36.56	29.28	.51	.48	.53	.50	.46	.53	.40	.44	.48	.46	.46	.51
		LST	124	57.03	36.22	.42	.57	.55	.41	.57	.55	.29	.51	.48	.34	.52	.49
		TOT	124	93.60	54.56	.55	.63	.65	.54	.63	.65	.41	.57	.58	.47	.59	.60
	A-3	MST	124	.44	.50	.46	.43	.48	.47	.41	.48	.39	.39	.45	.42	.42	.46
		LST	124	.68	.47	.43	.56	.54	.41	.56	.55	.30	.50	.48	.34	.51	.48
		TOT	124	1.11	.79	.55	.61	.63	.54	.59	.63	.42	.55	.57	.47	.57	.58
	A-4	MST	124	.38	.59	.54	.51	.57	.52	.49	.55	.42	.46	.51	.49	.49	.54
		LST	124	.61	.61	.37	.54	.51	.36	.54	.51	.24	.48	.44	.28	.49	.44
		TOT	124	.99	1.00	.54	.63	.64	.52	.62	.63	.39	.56	.56	.46	.58	.59
47	A-1	MST	123	3.66	17.15	.18	.23	.23	.18	.24	.23	.11	.22	.20	.12	.21	.19
		LST	123	3.20	17.04	.03	.07	.06	.01	.08	.05	-0.01	.01	.00	-0.01	.01	.00
		TOT	123	6.85	25.88	.14	.20	.19	.12	.21	.19	.06	.15	.13	.08	.14	.13
	A-2	MST	123	23.15	10.81	.27	.28	.29	.26	.28	.30	.19	.27	.27	.20	.26	.26
		LST	123	21.60	10.22	.06	.08	.07	.02	.09	.07	-0.01	-0.01	-0.01	.01	.01	.01
		TOT	123	44.76	15.32	.23	.24	.26	.20	.26	.26	.13	.18	.18	.15	.19	.19
	A-3	MST	123	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		LST	123	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		TOT	123	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
	A-4	MST	123	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		LST	123	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		TOT	123	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
48	A-1	MST	122	14.36	28.61	.21	.10	.16	.25	.07	.16	.18	.09	.15	.14	.09	.12
		LST	122	8.03	22.01	.17	.07	.12	.17	.04	.10	.13	.03	.09	.15	.03	.09
		TOT	122	22.39	39.75	.24	.11	.18	.27	.07	.18	.20	.08	.15	.18	.09	.14
	A-2	MST	122	31.16	22.66	.19	.09	.14	.23	.07	.15	.16	.08	.13	.12	.09	.11
		LST	122	23.31	9.81	.24	.17	.22	.25	.15	.21	.18	.12	.17	.21	.13	.18
		TOT	122	54.47	25.02	.26	.15	.21	.30	.12	.22	.22	.12	.19	.19	.13	.17
	A-3	MST	122	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		LST	122	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		TOT	122	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
	A-4	MST	122	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		LST	122	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		TOT	122	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00

Appendix D

Leadership Problems Inventory
Pilot Test Samples

Note: This appendix contains: Rank difference data for the NCO pilot sample and rank difference data for the undergraduate student pilot sample. Rank difference data are provided for each option of each item. The rank differences were summed across options to yield an item-level rank difference score. We examined these item-level rank differences when making decisions about which items to retain for the validation version of the LPI. Items which had the largest rank difference scores were not retained.

For purposes of confidentiality, values in the "key rank" and "key - aggregate" rank difference" columns were deleted from the following pages.

Leadership Problems Inventory
Item Level Statistics from the NCO Pilot Study

Item #1; Form A: (N=125)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	1.90		1	
B	2.22		2	
C	3.56		4	
D	3.54		3	
E	3.79		5	
Total				(4)

Item #2; Form A: (N=118)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.95		3	
B	2.70		2	
C	2.15		1	
D	3.04		4	
E	4.15		5	
Total				(4)

Item #3; Form A: (N=121)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.76		2	
B	2.94		3	
C	2.28		1	
D	3.94		5	
E	3.07		4	
Total				(6)

Note. M Ranking = mean of the rankings assigned by participants in this study. Key Rank = Rank determined from USASMA keying study. Aggregate Rank = Rank of option when M Ranking are placed in ascending order. Key-Aggregate Rank Difference = Result when the Aggregate Rank is subtracted from the Key Rank.

Leadership Problems Inventory
Item Level Statistics from the NCO Pilot Study

Item #4; Form A: (N=122)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.24		1	
B	2.83		2	
C	3.98		5	
D	2.83		2	
E	3.12		4	
Total				(7)

Item #5; Form A: (N=125)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	3.86		5	
B	2.70		2	
C	3.01		3	
D	3.80		4	
E	1.64		1	
Total				(2)

Item #6; Form A: (N=121)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	3.53		4	
B	1.70		1	
C	3.29		3	
D	4.15		5	
E	2.33		2	
Total				(4)

Item #7; Form A: (N=119)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.34		2	
B	3.33		3	
C	1.85		1	
D	4.09		5	
E	3.39		4	
Total				(2)

Leadership Problems Inventory
Item Level Statistics from the NCO Pilot Study

Item #8; Form A: (N=120)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	1.86		1	
B	1.88		2	
C	3.60		4	
D	3.41		3	
E	4.25		5	
Total				(2)

Item #9; Form A: (N=121)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	3.64		4	
B	2.10		1	
C	2.18		2	
D	3.07		3	
E	4.00		5	
Total				(0)

Item #10; Form A: (N=118)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.53		1	
B	2.82		2	
C	3.64		5	
D	3.18		4	
E	2.83		3	
Total				(7)

Item #11; Form A: (N=119)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.94		3	
B	3.00		4	
C	2.76		2	
D	2.34		1	
E	3.96		5	
Total				(2)

Leadership Problems Inventory
Item Level Statistics from the NCO Pilot Study

Item #12; Form A: (N=120)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	3.43		4	
B	2.81		2	
C	2.22		1	
D	2.86		3	
E	3.67		5	
Total				(4)

Item #13; Form A: (N=118)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.82		2	
B	2.89		4	
C	3.71		5	
D	2.86		3	
E	2.71		1	
Total				(4)

Item #14; Form A: (N=121)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	3.88		5	
B	2.44		2	
C	3.37		3	
D	1.85		1	
E	3.46		4	
Total				(4)

Item #15; Form A: (N=120)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	1.98		1	
B	3.55		5	
C	3.08		3	
D	3.35		4	
E	3.04		2	
Total				(2)

Leadership Problems Inventory
Item Level Statistics from the NCO Pilot Study

Item #16; Form A: (N=122)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	3.16		2	
B	3.17		3	
C	3.22		4	
D	3.38		5	
E	2.07		1	
Total				(8)

Item #17; Form A: (N=120)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.08		1	
B	2.33		2	
C	2.85		3	
D	3.49		4	
E	4.25		5	
Total				(4)

Item #18; Form A: (N=117)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.12		2	
B	1.94		1	
C	3.30		3	
D	4.01		5	
E	3.63		4	
Total				(2)

Item #19; Form A: (N=121)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.07		2	
B	2.02		1	
C	3.73		4	
D	2.87		3	
E	4.31		5	
Total				(2)

Leadership Problems Inventory
Item Level Statistics from the NCO Pilot Study

Item #20; Form A: (N=119)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	3.34		3	
B	1.77		1	
C	3.75		4	
D	4.03		5	
E	2.10		2	
Total				(4)

Item #21; Form A: (N=121)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	3.56		4	
B	2.78		3	
C	2.02		1	0
D	2.77		2	2
E	3.87		5	
Total				(6)

Item #22; Form A: (N=118)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.75		3	
B	2.39		1	
C	2.69		2	
D	3.36		4	
E	3.81		5	
Total				(10)

Item #23; Form A: (N=119)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	3.65		4	
B	2.79		2	
C	2.87		3	
D	3.70		5	
E	1.99		1	
Total				(2)

Leadership Problems Inventory
Item Level Statistics from the NCO Pilot Study

Item #24; Form A: (N=121)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	3.69		5	
B	2.48		2	
C	2.15		1	
D	3.31		3	
E	3.37		4	
Total				(4)

Item #1; Form B: (N=121)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.39		1	
B	2.50		2	
C	3.63		5	
D	3.02		3	
E	3.46		4	
Total				(6)

Item #2; Form B: (N=120)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	3.39		4	
B	2.42		2	
C	2.08		1	
D	4.38		5	
E	2.72		3	
Total				(4)

Item #3; Form B: (N=121)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.42		1	
B	2.70		2	
C	3.16		4	
D	3.70		5	
E	3.02		3	
Total				(2)

Leadership Problems Inventory
Item Level Statistics from the NCO Pilot Study

Item #4; Form B: (N=116)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.48		2	
B	2.50		3	
C	3.56		4	
D	2.46		1	
E	4.00		5	
Total				(2)

Item #5; Form B: (N=118)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	1.78		1	
B	3.77		4	
C	3.25		3	
D	3.95		5	
E	2.25		2	
Total				(4)

Item #6; Form B: (N=119)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.37		2	
B	1.96		1	
C	3.21		3	
D	3.33		4	
E	4.13		5	
Total				(0)

Item #7; Form B: (N=121)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	1.93		1	
B	2.69		3	
C	2.48		2	
D	3.69		4	
E	4.20		5	
Total				(0)

Leadership Problems Inventory
Item Level Statistics from the NCO Pilot Study

Item #8; Form B: (N=123)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.12		2	
B	2.08		1	
C	2.99		3	
D	3.65		4	
E	4.15		5	
Total				(4)

Item #9; Form B: (N=119)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	3.13		3	
B	1.76		1	
C	2.48		2	
D	3.47		4	
E	4.16		5	
Total				(8)

Item #10; Form B: (N=119)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	3.48		4	
B	2.83		2	
C	3.66		5	
D	2.93		3	
E	2.09		1	
Total				(6)

Item #11; Form B: (N=121)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.80		2	
B	3.20		4	
C	3.13		3	
D	3.55		5	
E	2.31		1	
Total				(2)

Leadership Problems Inventory
Item Level Statistics from the NCO Pilot Study

Item #12; Form B: (N=120)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	3.83		5	
B	2.94		3	
C	2.13		1	
D	3.67		4	
E	2.42		2	
Total				(2)

Item #13; Form B: (N=121)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.91		3	
B	2.87		2	
C	2.98		4	
D	3.83		5	
E	2.42		1	
Total				(10)

Item #14; Form B: (N=121)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.23		1	
B	2.79		3	
C	3.52		4	
D	2.60		2	
E	3.85		5	
Total				(4)

Item #15; Form B: (N=123)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.51		1	
B	3.24		4	
C	2.87		3	
D	3.71		5	
E	2.67		2	
Total				(4)

Leadership Problems Inventory
Item Level Statistics from the NCO Pilot Study

Item #16; Form B: (N=121)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.59		2	
B	2.47		1	
C	2.94		3	
D	3.79		5	
E	3.21		4	
Total				(6)

Item #17; Form B: (N=123)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	3.23		4	
B	2.70		2	
C	2.23		1	
D	3.75		5	
E	3.10		3	
Total				(2)

Item #18; Form B: (N=116)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	1.96		1	
B	2.85		2	
C	3.41		4	
D	3.66		5	
E	3.13		3	
Total				(2)

Item #19; Form B: (N=118)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.01		1	
B	3.04		3	
C	3.53		4	
D	2.51		2	
E	3.91		5	
Total				(4)

Leadership Problems Inventory
Item Level Statistics from the NCO Pilot Study

Item #20; Form B: (N=122)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.23		1	
B	3.28		4	
C	3.24		3	
D	3.90		5	
E	2.35		2	
Total				(2)

Item #21; Form B: (N=120)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.92		3	
B	3.46		4	
C	2.55		2	
D	2.18		1	
E	3.88		5	
Total				(2)

Item #22; Form B: (N=122)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.44		1	
B	2.66		2	
C	3.39		4	
D	2.73		3	
E	3.79		5	
Total				(8)

Item #23; Form B: (N=121)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	3.08		4	
B	2.90		2	
C	2.94		3	
D	2.16		1	
E	3.92		5	
Total				(6)

Leadership Problems Inventory
Item Level Statistics from the NCO Pilot Study

Item #24; Form B: (N=120)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.87		2	
B	3.52		4	
C	2.03		1	
D	3.05		3	
E	3.53		5	
Total				(2)

Appendix D (continued)

Leadership Problems Inventory
Item Level Statistics from the Undergraduate Student Study

Leadership Problems Inventory
Item Level Statistics from the Undergraduate Student Study

Item #1; Form A: (N=58)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.14		1	
B	2.90		3	
C	2.55		2	
D	2.95		4	
E	4.47		5	
Total				(4)

Item #2; Form A: (N=54)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	3.37		4	
B	2.35		2	
C	1.93		1	
D	3.20		3	
E	4.15		5	
Total				(4)

Item #3; Form A: (N=58)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	3.48		4	
B	2.38		2	
C	2.28		1	
D	3.76		5	
E	3.10		3	
Total				(6)

Note. M Ranking = mean of the rankings assigned by participants in this study. Key Rank = Rank determined from USASMA keying study. Aggregate Rank = Rank of option when M Ranking are placed in ascending order. Key-Aggregate Rank Difference = Result when the Aggregate Rank is subtracted from the Key Rank.

Appendix D (Continued)
Leadership Problems Inventory
Item Level Statistics from the Undergraduate Student Study

Item #4; Form A: (Total N=57)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.35		2	
B	3.93		5	
C	3.30		4	
D	2.28		1	
E	3.14		3	
Total				(4)

Item #5; Form A: (N=57)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.32		1	
B	2.40		2	
C	3.14		3	
D	4.00		5	
E	3.14		3	
Total				(9)

Item #6; Form A: (N=56)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	3.30		4	
B	1.98		1	
C	3.27		3	
D	2.79		2	
E	3.66		5	
Total				(8)

Item #7; Form A: (N=57)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.86		3	
B	3.89		5	
C	1.74		1	
D	2.84		2	
E	3.67		4	
Total				(6)

Item #8; Form A: (N=58)

Appendix D (Continued)
Leadership Problems Inventory
Item Level Statistics from the Undergraduate Student Study

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	1.83		1	
B	3.84		5	
C	2.97		2	
D	3.07		3	
E	3.29		4	
Total				(8)

Item #9; Form A: (N=57)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.86		4	
B	2.60		2	
C	2.30		1	
D	2.68		3	
E	4.56		5	
Total				(2)

Item #10; Form A: (N=56)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.79		2	
B	2.43		1	
C	3.61		5	
D	3.11		4	
E	3.07		3	(9)

Item #11; Form A: (N=58)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	3.38		4	
B	2.02		1	
C	2.22		2	
D	4.03		5	
E	3.34		3	
Total				(8)

Appendix D (Continued)
Leadership Problems Inventory
Item Level Statistics from the Undergraduate Student Study

Item #12; Form A: (N=57)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.98		3	
B	2.54		2	
C	2.16		1	
D	4.11		5	
E	3.21		4	
Total				(6)

Item #13; Form A: (N=57)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	3.25		4	
B	2.63		2	
C	3.19		3	
D	4.37		5	
E	1.56		1	
Total				(12)

Item #14; Form A: (N=57)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	3.53		5	
B	2.70		2	
C	2.86		3	
D	2.61		1	
E	3.30		4	
Total				(4)

Item #15; Form A: (N=58)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	1.57		1	
B	4.40		5	
C	3.03		3	
D	2.78		2	
E	3.22		4	
Total				(4)

Appendix D (Continued)
Leadership Problems Inventory
Item Level Statistics from the Undergraduate Student Study

Item #16; Form A: (N=55)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	4.33		5	
B	2.93		3	
C	2.24		1	
D	3.16		4	
E	2.35		2	
Total				(10)

Item #17; Form A: (N=56)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.25		1	
B	2.80		2	
C	2.95		4	
D	4.20		5	
E	2.80		2	
Total				(9)

Item #18; Form A: (N=54)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.09		1	
B	2.59		2	
C	2.93		3	
D	3.76		5	
E	3.63		4	
Total				(4)

Item #19; Form A: (N=56)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	1.34		1	
B	3.07		2	
C	3.41		3	
D	3.41		3	
E	3.77		5	
Total				(5)

Appendix D (Continued)
Leadership Problems Inventory
Item Level Statistics from the Undergraduate Student Study

Item #20; Form A: (N=55)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	3.73		5	
B	2.85		3	
C	2.80		2	
D	3.15		4	
E	2.47		1	
Total				(6)

Item #21; Form A: (N=57)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.89		2	
B	3.18		3	
C	2.09		1	
D	3.61		5	
E	3.23		4	
Total				(6)

Item #22; Form A: (N=54)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	1.67		1	
B	3.50		4	
C	3.61		5	
D	3.19		3	
E	3.04		2	
Total				(6)

Item #23; Form A: (N=57)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	3.54		5	
B	2.18		1	
C	3.37		3	
D	3.39		4	
E	2.53		2	
Total				(6)

Appendix D (Continued)
Leadership Problems Inventory
Item Level Statistics from the Undergraduate Student Study

Item #24; Form A: (N=57)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	4.00		5	
B	3.32		4	
C	1.96		1	
D	2.93		3	
E	2.79		2	
Total				(2)

Item #1; Form B: (N=60)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.50		1	
B	3.02		4	
C	3.98		5	
D	2.85		3	
E	2.65		2	
Total				(8)

Item #2; Form B: (N=58)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.93		3	
B	2.62		2	
C	2.22		1	
D	3.88		5	
E	3.34		4	
Total				(4)

Item #3; Form B: (N=58)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.26		1	
B	3.21		4	
C	2.97		2	
D	3.59		5	
E	2.98		3	
Total				(4)

Appendix D (Continued)
Leadership Problems Inventory
Item Level Statistics from the Undergraduate Student Study

Item #4; Form B: (N=59)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.00		1	
B	2.53		2	
C	3.00		4	
D	2.97		3	
E	4.51		5	
Total				(4)

Item #5; Form B: (N=58)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	3.97		5	
B	2.91		3	
C	2.81		2	
D	3.55		4	
E	1.76		1	
Total				(6)

Item #6; Form B: (N=59)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.98		2	
B	1.86		1	
C	3.15		3	
D	3.51		5	
E	3.49		4	
Total				(2)

Item #7; Form B: (N=60)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	3.63		5	
B	3.33		4	
C	2.28		1	
D	3.22		3	
E	2.53		2	
Total				(10)

Item #8; Form B: (N=59)

Appendix D (Continued)
Leadership Problems Inventory
Item Level Statistics from the Undergraduate Student Study

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	1.98		1	
B	2.20		2	
C	3.37		3	
D	3.76		5	
E	3.68		4	
Total				(4)

Item #9; Form B: (N=60)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	3.52		5	
B	1.82		1	
C	3.27		4	
D	3.25		3	
E	3.15		2	
Total				(0)

Item #10; Form B: (N=60)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.67		2	
B	3.55		5	
C	3.28		4	
D	2.23		1	
E	3.27		3	
Total				(4)

Item #11; Form B: (N=58)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.74		2	
B	2.76		3	
C	3.10		4	
D	3.81		5	
E	2.59		1	
Total				(0)

Appendix D (Continued)
Leadership Problems Inventory
Item Level Statistics from the Undergraduate Student Study

Item #12; Form B: (N=58)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	4.07		5	
B	2.98		3	
C	3.47		4	
D	1.74		1	
E	2.74		2	
Total				(6)

Item #13; Form B: (N=59)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	3.92		5	
B	2.80		3	
C	2.69		1	
D	2.83		4	
E	2.76		2	
Total				(8)

Item #14; Form B: (N=59)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	1.78		1	
B	3.02		3	
C	3.24		4	
D	2.56		2	
E	4.41		5	
Total				(4)

Item #15; Form B: (N=57)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.81		2	
B	3.91		5	
C	3.12		3	
D	3.39		4	
E	1.77		1	
Total				(0)

Appendix D (Continued)
Leadership Problems Inventory
Item Level Statistics from the Undergraduate Student Study

Item #16; Form B: (N=59)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.73		2	
B	2.47		1	
C	2.90		3	
D	2.92		4	
E	3.98		5	
Total				(4)

Item #17; Form B: (N=60)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.40		1	
B	2.58		2	
C	3.53		4	
D	2.87		3	
E	3.62		5	
Total				(10)

Item #18; Form B: (N=57)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.61		2	
B	1.54		1	
C	2.95		3	
D	3.91		4	
E	3.98		5	
Total				(4)

Item #19; Form B: (N=60)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.87		3	
B	1.48		1	
C	3.50		4	
D	2.68		2	
E	4.47		5	
Total				(8)

Item #20; Form B: (N=60)

Appendix D (Continued)
Leadership Problems Inventory
Item Level Statistics from the Undergraduate Student Study

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.65		2	
B	2.77		3	
C	3.60		4	
D	3.75		5	
E	2.23		1	
Total				(2)

Item #21; Form B: (N=60)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.70		2	
B	2.82		3	
C	3.37		4	
D	1.95		1	
E	4.17		5	
Total				(4)

Item #22; Form B: (N=59)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.34		2	
B	2.27		1	
C	3.56		4	
D	2.75		3	
E	4.08		5	
Total				(8)

Item #23; Form B: (N=59)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	2.00		1	
B	3.15		3	
C	3.76		4	
D	2.25		2	
E	3.83		5	
Total				(2)

Appendix D (Continued)
Leadership Problems Inventory
Item Level Statistics from the Undergraduate Student Study

Item #24; Form B: (N=59)

Option	M Ranking	Key Rank	Aggregate Rank	Key-Aggregate Rank Difference
A	3.08		3	
B	3.61		5	
C	2.34		1	
D	2.81		2	
E	3.15		4	
Total				(6)